

Description

[0001] Heterologous surface display of proteins (Stahl and Uhlen, TIBTECH May 1997, 15, 185-192) on recombinant micro-organisms via the targeting and anchoring of heterologous proteins to the outer surface of host-cells such as yeast, fungi, mammalian and plant cells, and bacteria has been possible for several years. Display of heterologous proteins at the surface of these cells has taken many forms, varying from the expression of reactive groups such as antigenic determinants, heterologous enzymes, (single-chain) antibodies, polyhistidyl tags, peptides, and other compounds. Heterologous surface display has been applied as a tool for applied and fundamental research in microbiology, molecular biology, vaccinology and biotechnology, and several patent applications have been filed.

[0002] Yet another application of bacterial surface display has been the development of live-bacterial-vaccine delivery systems. The cell-surface display of heterologous antigenic determinants has been considered advantageous for the induction of antigen-specific immune responses when using live recombinant cells for immunisation. Another application has been the use of bacterial surface display in generating whole-cell bioadsorbents or biofilters for environmental purposes, microbiocatalysts, and diagnostic tools.

[0003] In general, one has used chimeric proteins consisting of an anchoring or targeting part specific and selective for the recombinant organism used and has combined this part with a part comprising a reactive group as described above. In this way, chimeric proteins are composed of at least two (poly)peptides of different genetic origin joined by a normal peptide bond. For example, in patent application WO 94/18830 relating to the isolation of compounds from complex mixtures and the preparation of immobilised ligands (bioadsorbents), a method has been claimed for obtaining such a ligand which comprises anchoring a binding protein in or at the exterior of the cell wall of a recombinant cell. Said binding protein is essentially a chimeric protein produced by said recombinant cell, and is composed of an N-terminal part, derived from for example an antibody, that is capable of binding to a specific compound joined with a C-terminal anchoring part, derived from an anchoring protein purposely selected for being functional in the specific cell chosen. In patent application WO 97/08553 a method has been claimed for the targeting of proteins selectively to the cell wall of *Staphylococcus* spp, using as anchoring proteins long stretches of at least 80-90 amino acid long amino acid cell wall-targeting signals derived from the lysostaphin gene or amidase gene of *Staphylococcus* which encode for proteins that selectively bind to *Staphylococcus* cell wall components.

[0004] Vaccine delivery or immunisation via attenuated bacterial vector strains expressing distinct antigenic determinants against a wide variety of diseases is now commonly being developed. Recently, mucosal (for example nasal or oral) vaccination using such vectors has received a great deal of attention. For example, both systemic and mucosal antibody responses against an antigenic determinant of the hornet venom were detected in mice orally colonised with a genetically engineered human oral commensal *Streptococcus gordonii* expressing said antigenic determinant on its surface (Medaglini et al., PNAS 1995, 2; 6868-6872). Also, a protective immune response could be elicited by oral delivery of a recombinant bacterial vaccine wherein tetanus toxin fragment C was expressed constitutively in *Lactococcus lactis* (Robinson et al., Nature Biotechnology 1997, 15; 653-657). Especially mucosal immunisation as a means of inducing IgG and secretory IgA antibodies directed against specific pathogens of mucosal surfaces is considered an effective route of vaccination. Immunogens expressed by bacterial vectors are presented in particulate form to the antigen-presenting cells (for example M-cells) of the immune system and should therefore be less likely to induce tolerance than soluble antigens. In addition, the existence of a common mucosal immune system permits immunisation on one specific mucosal surface to induce secretion of antigen-specific IgA, and other specific immune responses at distant mucosal sites. A drawback to this approach is the potential of the bacterial strain to cause inflammation and disease in itself, potentially leading to fever and bacteraemia. An alternative approach avoids the use of attenuated bacterial strains that may become pathogenic themselves by choosing recombinant commensal bacteria as vaccine carriers, such as *Streptococcus* spp. and *Lactococcus* spp.

[0005] However, a potential problem with such recombinant organisms is that they may colonise the mucosal surfaces, thereby generating a long term exposure to the target antigens expressed and released by these recombinant micro-organisms. Such long term exposure can cause immune tolerance. In addition, the mere fact alone that such organisms are genetically modified and contain recombinant nucleic acid is meeting considerable opposition from the (lay) public as a whole, stemming from a low level of general acceptance for products containing recombinant DNA or RNA. Similar objections exist against the use of (even attenuated) strains of a pathogenic nature or against proteins or parts of proteins derived from pathogenic strains. However, as explained above, present techniques of heterologous surface display of proteins in general entail the use of anchoring or targeting proteins that are specific and selective for a limited set of micro-organisms which in general are of recombinant or pathogenic nature, thereby greatly restricting their potential applications.

[0006] The invention provides substances and methods to anchor or attach said substances to a cell-wall or cell-wall component of a wide range of micro-organisms. A preferred embodiment of the invention provides substances and methods to attach said substances to non-recombinant micro-organisms. Said substances provided by the invention are not limited to (chimeric) proteins alone, but can be fully or only partly of a peptide nature, whereby a peptide part is

(covalently) joined to a non-peptide moiety. The invention provides a proteinaceous substance comprising at least one stretch of amino acids derived from a first micro-organism which substance is capable of attaching to a cell-wall of a second micro-organism. Said substance according to the invention is for example produced by a first micro-organism (for example a micro-organism from which the knowledge about the sequence of said stretch of amino acids originates, but another (recombinant) micro-organism can produce said substance as well). After its production said substance is harvested, optionally stored for future use, and then brought in contact with said second micro-organism, where it attaches to its cell-wall. Alternatively, said substance is produced synthetically, by using established peptide synthesis technology. A preferred embodiment of the invention provides a substance wherein said second micro-organism is a non-recombinant micro-organism. With a substance provided by the invention it is now possible to attach or anchor for example a heterologous or chimeric protein produced by a recombinant micro-organism to an innocuous non-recombinant micro-organism.

[0007] A preferred embodiment of the invention provides a proteinaceous substance wherein said stretch of amino acids has a sequence corresponding to a consensus sequence listed in figure 10, or wherein said stretch of amino acids (herein also called attaching peptide) has a sequence corresponding to a sequence selected from those listed in figure 11, or a homologous sequence derived from another species. The sequences listed in figure 11, and sequences homologous thereto, are found in a variety of species, both micro-organisms and higher organisms, an example of such a higher organism is *C. elegans*. Preferably, the attaching peptide is derived from any one of the proteins listed in figure 11, more preferably said attaching peptide comprises an amino acid sequence as shown in figure 10, or a sequence derived thereof. For example, the invention provides a proteinaceous substance wherein said attaching peptide is derived from the major peptidoglycan hydrolase of *Lactococcus lactis*.

[0008] Yet another preferred embodiment of the invention provides a proteinaceous substance wherein said second micro-organism is selected from any of the group of Gram-positive bacteria and Gram-negative bacteria. Examples are micro-organisms, such as *Bacillus subtilis*, *Clostridium beijerinckii*, *Lactobacillus plantarum*, *Lb. buchneri*, *Listeria innocua*, *Streptococcus thermophilus*, *Enterococcus faecalis*, *E. coli*, and others.

[0009] The invention provides a proteinaceous substance which is additionally comprising a reactive group. For example, the invention provides a proteinaceous substance comprising a reactive group such as an antigenic determinant, heterologous enzyme, (single-chain) antibody or fragment thereof, polyhistidyl tag, fluorescing protein, luciferase, binding protein or peptide, or another substance such as an antibiotic, hormone, non-peptide antigenic determinant, carbohydrate, fatty acid, aromatic substance and reporter molecule, and an anchoring or targeting protein or part thereof (herein also called attaching peptide) useful in heterologous surface display which is both broadly reactive with cell wall components of a broad range of micro-organisms.

[0010] For example, the invention provides a substance wherein said reactive group is a non-protein moiety, for example is selected from the group of antibiotics, hormones, aromatic substances and reporter molecules. Said substance is constructed by binding for example an antibiotic, such as penicillin or tetracycline, but various other antibiotics can be used, or a hormone, such as a steroid hormone, or any other compound to an attaching peptide provided by the invention. Such binding can be achieved by various techniques known in the art, and thereby can label or "flag" the attaching peptide. A preferred example is the binding of an attaching peptide to a reporter molecule such as FITC, or HRPO, whereby tools are generated that can be used in diagnostic assay whereby micro-organisms having peptidoglycan are detected. Similarly, an attaching peptide with an antibiotic bound thereto can be used *in vivo* by for example parenteral administration into the bloodstream of humans or animals or *in vitro* to bind to such micro-organisms having peptidoglycan, thereby increasing the concentration of antibiotic around said organism, which then gets killed by the antibiotic action.

[0011] The invention provides a substance wherein said reactive group is a protein moiety, for example selected from the group of antigenic determinants, enzymes, (single-chain) antibodies or fragments thereof, polyhistidyl tags, fluorescing proteins, binding proteins or peptides. For example, the invention provides a protein, which comprises as reactive group a protein or (poly)peptide. Also, the invention provides a nucleic acid molecule encoding a protein provided by the invention. Such a nucleic acid molecule (being single- or double stranded DNA, RNA or DNA/RNA) at least comprises nucleic acid sequences specifically encoding a attaching peptide as well as nucleic acid sequences specifically encoding the reactive group polypeptide, but can additionally also comprise other nucleic acid sequences, which for example encode a signal peptide, or comprise for example promoter and/or regulatory nucleic acid sequences. The invention also provides a vector comprising a nucleic acid molecule encoding a protein provided by the invention.

[0012] The invention provides a proteinaceous substance comprising a reactive group joined with or bound to at least one attaching peptide which comprises a stretch of amino acids corresponding to the consensus amino acid sequence listed in figure 10, said substance capable of attaching or anchoring or binding to a cell wall component of a micro-organism.

[0013] Corresponding to is defined as having an amino acid sequence homologous to the consensus amino acid sequence listed in figure 10, or having an amino acid sequence derived of the sequence listed in figure 10 which derived sequence comprises a functionally equivalent stretch of amino acids.

[0014] Preferably, the attaching peptide is derived from any one of the proteins listed in figure 11, or a protein having a repeat sequence related or homologous to the sequence listed in figure 10, more preferably said attaching peptide comprises an amino acid sequence as shown in figure 10, or a sequence derived thereof. Homology between the various amino acid sequences of related attaching peptides provided by the invention can for instance be determined by performing a homology search between amino acid sequences, such as for example can be found in a protein data-base, such as the SWISSPROT, PIR and Genbank databases, using a computer programme, such as the BLAST programme, that can determine homology between amino acid sequences. For example, the invention provides a proteinaceous substance wherein said attaching peptide is derived from the major peptidoglycan hydrolase of *Lactococcus lactis*. The invention provides a proteinaceous substance comprising a reactive compound wherein at least two stretches of amino acids, corresponding to an attaching peptide sequence, are located adjacent to each other, possibly separated by one or more amino acid residues. Said stretches or repeats can be separated by a short distance, for example 3-6 to 10-15 amino acids apart, or by a medium distance 15-100 amino acids apart, or by longer distances (>100 amino acid residues apart). Examples of such distances can be found in figure 11, but longer distances are also possible. The distances between said stretches or repeats can also be used for an (additional) reactive group, whereby a reactive group is inserted between repeats, thereby allowing an even better anchoring to a cell wall component. A preferred embodiment provided by the invention is a proteinaceous substance comprising a reactive group and at least one attaching peptide which comprises a stretch of amino acids having a sequence corresponding to the consensus amino acid sequence listed in figure 10, wherein said substance is capable of attaching to a cell wall component of a micro-organism, such as can be found among from any of the group of yeast, moulds, Gram-positive bacteria and Gram-negative bacteria. Examples are micro-organisms, such as *Bacillus subtilis*, *Clostridium beijerinckii*, *Lactobacillus plantarum*, *Lb. buchneri*, *Listeria innocua*, *Streptococcus thermophilus*, *Enterococcus faecalis*, *E. coli*, and others. A preferred embodiment provided by the invention is a proteinaceous substance which is capable of attaching to a cell wall component of a conventional (non-recombinant) micro-organism. In this embodiment, the invention provides for example non-recombinant organisms which displaying heterologous proteins, these may colonise the mucosal surfaces without causing problems such as immune tolerance, since they do not generate a long term exposure to the target antigens expressed. In addition, the mere fact alone that such organisms provided by the invention are not genetically modified and do not contain recombinant nucleic acid will alleviate the opposition from the (lay) public as a whole against recombinant micro-organisms, which is stemming from a low level of general acceptance for products containing recombinant DNA or RNA. Similar objections that exist against the use of (even attenuated) strains of a pathogenic nature or against proteins or parts of proteins derived from pathogenic strains are now also overcome by the invention, in that it is now possible to attach a proteinaceous substance to a non-recombinant, non-pathogenic micro-organism, such as *L. lactis* which is generally considered as safe. The invention provides a proteinaceous substance comprising a reactive group such as an antigenic determinant, (heterologous) enzyme, (single-chain) antibody or fragment thereof, polyhistidyl tag, fluorescing protein, luciferase, binding protein or peptide, or another compound such as an antibiotic, hormone, non-peptide antigenic determinant, carbohydrate, fatty acid, aromatic compound and reporter molecule, and an anchoring or targeting protein or part thereof (herein also called attaching peptide) useful in heterologous surface display which is both broadly reactive with cell wall components of a broad range of micro-organisms. Said attaching peptide is preferably derived from a micro-organism which is generally recognised as safe (G.R.A.S.), thereby greatly enhancing the potential of applications of the heterologous surface display technique. *Lactococcus lactis* is a non-pathogenic, non-invasive, and non-colonising Gram-positive bacterium which is not adapted for growth in body or even the gut; it does not belong to the commensal species of lactic acid bacteria. *L. lactis* has a history of safe use of several thousand years. The major cell wall hydrolase AcmA of the Gram-positive bacterium *Lactococcus lactis* subsp. *cremoris* MG1363 is an N-acetylmuramidase which is required for cell separation and is responsible for cell lysis during stationary phase. The protein consists of three separate domains (Fig. 9, Buist et al., J. Bacteriol. (1995) 177:1554-1563) of which the first 57 amino acids of the N-terminal domain encompasses the signal peptide needed for secretion. This domain is followed by the active site domain running from the Ala at position 58 to Ser-218. The active site domain was overproduced in and purified from *Escherichia coli* as a thioredoxin fusion protein. The AcmA part was released by proteolytic cleavage with enterokinase and shown to be active *in vitro*. Three homologous repeated regions (or stretches of amino acids) of 35-55, more often 40-50 amino acid residues are present in the C-terminus of for example AcmA which are separated by non-homologous sequences (Fig.10). The repeat sequences of AcmA (cA) can be deleted and additional repeat sequences could be added without impairing cell wall hydrolysing activity *in vitro*. The AcmA deletion derivatives lacking one or two repeat sequences and the protein containing at least one additional repeat were able to bind to lactococcal cells when added from the outside. The derivative lacking all three repeats did not bind to the cells nor did the purified active site domain. The invention provides an attaching peptide that comprise at least one repeat sequence as shown in figure 10 or a sequence that is similar to the sequence of figure 10, similar being defined as comprising at least a part of a consensus sequence as shown in figure 11. Also, attaching peptides are provided by the invention which are comprising amino acid sequences that are derived from a sequence as shown in figure 11. Derived herein meaning among others by comparison with heterologous sequences whereby a consensus

sequence is obtained, or derived via conventional amino acid substitutions whereby amino acids are being substituted by like amino acids, or derived via substitutions whereby functional amino acids are being replaced by functionally alike or better amino acids identified by methods such as PEPSCAN techniques or replacement mapping. The invention provides a proteinaceous substance comprising a reactive group and at least one attaching peptide which comprises a stretch of amino acids having a sequence corresponding to at least a part of the consensus amino acid sequence provided in figure 10. Repeats similar to those in AcmA were for example shown to be present in various cell wall hydrolases and other (secreted) proteins of Gram-positive and Gram-negative bacteria and other micro-organisms and constitute a general cell wall binding domain in these proteins. An attaching peptide comprising at least one AcmA repeat or an amino acid sequence similar to the AcmA repeat provided by the invention represents a general and broadly reactive tool to bind or attach reactive groups such as antigenic determinants, enzymes, antibodies, proteins or peptides to cell walls of micro-organisms. Said repeat comprises a peptide composed of a stretch of amino acids having a sequence corresponding to at least a part of the consensus amino acid sequence provided in figure 10. Furthermore, we also demonstrated that an attaching peptide provided by the invention bound or attached to cells of other, e.g. non-recombinant micro-organisms, such as *Bacillus subtilis*, *Clostridium beijerinckii*, *Lactobacillus plantarum*, *Lb. buchneri*, *Listeria innocua*, *Streptococcus thermophilus*, *Enterococcus faecalis*, *E. coli*, and others. Binding of the attaching peptide and reactive group joined therewith, as provided by the invention is stable at pH values ranging from 2-10. Moreover, the attaching peptide provided by the invention is, when attached to the cell wall, protected against proteolytic degradation. One embodiment of the invention is a protein wherein the attaching peptide is derived from any of the proteins listed in figure 11. An example of such an attaching peptide is provided in the experimental part of this description wherein an attaching peptide having a sequence as shown in figure 10, or a sequence similar thereto, is used. Furthermore, the invention provides a protein, which comprises as reactive group a protein or (poly)peptide. Also, the invention provides a nucleic acid molecule encoding a protein provided by the invention. Such a nucleic acid molecule (being single- or double stranded DNA, RNA or DNA/RNA) at least comprises nucleic acid sequences specifically encoding an attaching peptide as well as nucleic acid sequences specifically encoding the reactive group polypeptide, but can additionally also comprise other nucleic acid sequences, which for example encode a signal peptide, or comprise for example promoter and/or regulatory nucleic acid sequences. The invention also provides a vector comprising a nucleic acid molecule encoding a protein provided by the invention. Such a vector can for example be a plasmid, phage, or virus, and can now be constructed using a nucleic acid provided by the invention and routine skills of the art. Examples of such a vector can be found in the experimental part of the description, other examples can e.g. be a baculovirus vector, or comparable vector viruses through which a protein provided by the invention can be expressed or produced in (insect) cells. The invention also provides a host cell or expression system comprising a nucleic acid molecule according to the invention or a vector according to the invention. Such a host cell expressing a protein is in itself provided by the invention as a micro-organism to which a protein provided by the invention is attached. Such a host cell or expression system can for example be a Gram-positive or Gram-negative bacterium, or a yeast cell or insect cell or plant- or mammalian cell, or even a cell-free expression system such as a reticulocyte lysate, and can now be constructed or obtained using a nucleic acid or vector provided by the invention and routine skills of the art. Examples of such a host cell or expression system can be found in the experimental part of the description, other examples can be obtained using a nucleic acid or vector provided by the invention and routine skills of the art.

[0015] The invention provides a method for attaching a substance to the cell wall of a micro-organism comprising the use of an attaching peptide which comprises a stretch of amino acids having a sequence corresponding to at least a part of the consensus amino acid sequence provided in figure 10. An example of the method provided by the invention is anchoring of recombinant poly(peptides), being (chimeric) proteins fused to the cell wall anchoring repeats of AcmA of *Lactococcus lactis* MG1363, to the cell wall of (Gram-positive) bacteria. The recombinant proteins are obtained by the expression of DNA sequences encoding these recombinant (poly)peptides in a suitable production strain (e.g. *E. coli* or *L. lactis*) and subsequent purification of the expression products. The recombinant proteins are then mixed, either *in vitro* or *in vivo*, with a non-recombinant target bacterium to obtain binding to the cell wall. Another example of the method provided by the invention is anchoring of recombinant poly(peptides), being (chimeric) proteins fused to the cell wall anchoring repeats of AcmA of *Lactococcus lactis*, to the cell wall of said recombinant *Lactococcus lactis* which produces the protein itself. In a preferred embodiment of the method provided by the invention the binding of (purified) proteins to bacterial cells upon addition from the outside, the method is an excellent tool to anchor recombinant proteins or other substances to non-recombinant bacterial cells.

[0016] A preferred method according to the invention comprises the use of an attaching peptide which is derived from the major peptidoglycan hydrolase of *Lactococcus lactis*. Another method according to the invention is provided wherein said substance is a (poly)peptide or a protein, for example being part of a protein provided by the invention. Post-translational modifications occurring to such a (poly)peptide or protein are inherent to the host cell or expression system used, a post-translationally modified protein as provided by the invention is therefore also provided. However, yet another method according to the invention is provided wherein said compound is selected from the group composed of antibiotics, hormones, antigenic determinants, carbohydrate chains, fatty acids, aromatic compounds and reporter

molecules. Said substance is constructed by binding for example an antibiotic, such as penicillin or tetracycline, but various other antibiotics can be used, or a hormone, such as a steroid hormone, or any other compound to an attaching peptide provided by the invention. Such binding can be achieved by various techniques known in the art, and thereby can label or "flag" the attaching peptide. A preferred example is the binding of an attaching peptide to a reporter molecule such as FITC, or HRP, whereby tools are generated that can be used in diagnostic assay whereby micro-organisms having peptidoglycan are detected. Similarly, an attaching peptide with an antibiotic bound thereto can be used *in vivo* by for example parenteral administration into the bloodstream of humans or animals or *in vitro* to bind to such micro-organisms having peptidoglycan, thereby increasing the concentration of antibiotic around said organism, which then can get killed by the antibiotic action. Said micro-organism is preferably selected from any of the group of yeast, moulds, Gram-positive bacteria and Gram-negative bacteria. For example, the experimental part of this description describes mixing of β -lactamase::cA fusion protein with lactococcal cells which resulted in binding to the cells whereas this was not the case when mature β -lactamase not joined with an attachment protein was added. Also, fusion of β -lactamase of *E. coli* and α -amylase of *Bacillus licheniformis* to the attaching peptide provided by the invention and subsequent production of these fusion proteins resulted in active, secreted proteins which were located (attached) in *L. lactis* cell walls. Binding of AcmA and the β -lactamase::cA fusion protein was also demonstrated to isolated lactococcal cell walls and SDS-washed cell walls (the major part of this fraction is peptidoglycan).

[0017] Anchoring of recombinant proteins to non-recombinant micro-organisms such as lactococci (or other bacteria) or fungi, is especially attractive if the use of recombinant bacteria is not desired, e.g. in food processes or as pharmaceuticals for medical use such as in vaccines or in anti-bacterial therapy. The invention provides for example vaccine delivery or immunisation via micro-organisms provided by the invention which are labelled with distinct antigenic determinants, which may be directed against a wide variety of diseases. A protective immune response can for example be elicited by oral delivery of a bacterial vaccine provided by the invention wherein tetanus toxin fragment C is attached via a protein provided by the invention to a non-recombinant *Lactococcus lactis*. Such immunogens expressed by micro-organisms provided by the invention are presented in particulate form to the antigen-presenting cells (for example M-cells) of the immune system and are therefore less likely to induce tolerance than soluble antigens. In addition, the existence of a common mucosal immune system permits immunisation on one specific mucosal surface to induce secretion of antigen-specific IgA, and other specific immune responses at distant mucosal sites. The invention solves the drawback of earlier bacterial vaccines whereby the potential to flourish on mucosal surfaces of the (attenuated or recombinant) bacterial strain used can cause problems such as inflammation and disease in itself, potentially leading to fever and bacteraemia, or to the induction of immune tolerance. Also, the invention avoids the potential risks that are involved when using recombinant DNA containing bacterial vectors for vaccination. In yet another possible vaccine and vaccine use provided by the invention, certain (killed) micro-organisms with adjuvant properties (such as the mycobacteria used in BCG) are labelled or loaded with a protein or substance composed of an antigenic determinant and an attaching peptide. These micro-organisms then function as adjuvant, thereby greatly enhancing the immune response directed against the specific antigenic determinant. Yet another use provided by the invention comprises anchoring proteins from the outside to a micro-organism which provides a means to present proteins or peptides which normally can not be overexpressed (and/or secreted) by said micro-organism. For example, the subunit B of cholera toxin (CTB) can be overproduced in *E. coli* but expression in *L. lactis* has been unsuccessful until now. The adjuvant activity of CTB in experimental recombinant vaccines is well documented and the ability of CTB or part thereof to bind to GM1 ganglioside on eucaryotic cell surfaces is of interest with respect to the use of *L. lactis* (or other Gram-positives) in vaccines which specifically require targeting to mucosal surfaces. Yet another medical use provided by the invention is the addition of (purified) antigen::cA fusion proteins *in vivo* by parenteral administration into the bloodstream of humans or animals to combat bacterial infections. In this case the antigen::cA fusion protein is used as a "flag" for the immune system. The antigenic determinant of a protein provided by the invention being a subunit of a vaccine regularly used for the immunisation of humans (preferably children) or animals, e.g. a subunit of the Rubella, Pertussis, Poliomyelitis, tetanus or measles vaccine. After delivery in the bloodstream, the "flag" will bind through the AcmA repeats to the pathogenic bacterium present in the blood. A "flag" protein provided by the invention will then activate a memory response, i.e. the response to the antigenic determinant present in the protein. The antibodies thus produced recognise the "flag" labelled bacteria, which will then be neutralised by the immune system. In this way the protein is used to stimulate a pre-existing (memory) immune response, non-related to the bacterial infection, to clear bacterial infections from the system. Yet another use (which alternatively may be considered medical use or food use) provided by the invention is the use wherein a protein provided by the invention has the ability to bind to cells, such as mucosal cells, e.g. of the gut. The reactive group of such a protein is in such a case for example partly or wholly derived from a fimbriae protein or another gut attachment protein, as for example present in various *E. coli* strains. Micro-organisms to which such a protein is attached will specifically home or bind to certain areas of the gut, a property which for example is beneficial for certain bacterial strains (i.e. lactococcal or lactobacillar strains) used as a probiotic. In another food or use of food provided by the invention, the protein or substance provided by the invention is a composed of a food additive (such as an enzyme or flavour compound) which affects quality, flavour, shelf-life, food value or texture, joined with an attaching peptide, and

subsequently attached or anchored to a micro-organism which is then mixed with the foodstuff. The anchoring of such proteins to a bacterial carrier offers the additional advantage that the additive can be targeted to a solid (bacteria-containing) matrix (e.g. curd) in a process for the preparation of food, e.g. cheese or tofu. Yet another use of a proteinaceous substance or micro-organism provided by the invention is the use of bacterial surface display in generating whole-cell bioadsorbents or biofilters for environmental purposes, biotransformations, and diagnostic tools

[0018] The invention is further explained in the experimental part which can not be seen as limiting the invention.

Experimental part

Introduction

[0019] The major autolysin AcmA of *Lactococcus lactis* subsp. *cremoris* MG1363 is an *N*-acetylmuramidase which is required for cell separation and is responsible for cell lysis during stationary phase (5, 6). The 40.3-kDa secreted mature protein produces a number of activity bands in a zymogram of the supernatant of a lactococcal culture. Bands as small as that corresponding to a protein of 29 kDa were detected. As no clearing bands are produced by an *L. lactis* acmA deletion mutant, all bands represent products of AcmA (6). From experimental data and homology studies we inferred that AcmA likely consists of three domains: a signal sequence followed by an active site domain and a C-terminal region containing three highly homologous repeats of approximately 45 amino acids which are involved in cell wall binding. As the smallest active protein is 29 kDa, it was suggested that the protein undergoes proteolytic breakdown in the C-terminal portion (5, 6).

[0020] Cell wall hydrolases of various bacteria and bacteriophages contain repeats similar to those present in AcmA (4, 9, 10, 17). Partially purified muramidase-2 of *Enterococcus hirae*, a protein similar to AcmA, containing 6 similar repeats, binds to peptidoglycan fragments of the strain (11). The p60 protein of *Listeria monocytogenes* contains two such repeats and was shown to be associated with the cell surface (24). However, which parts of these enzymes contained the binding capacity was not assessed in any of these studies.

[0021] Nearly all cell wall hydrolases examined so far seem to consist of a catalytic domain and usually, although not always, a domain containing a number of specific amino acid repeats. In several studies it has been shown that only a part of some of the cell wall hydrolases is required for enzymatic activity (13, 14, 17, 19, 22, 34). Rashid *et al.* reported the cloning of the gene encoding a 90-kDa glucosaminidase of *Bacillus subtilis* of which the C-terminus shows significant similarity with the glucosaminidase domain of the *S. aureus* autolysin (23). The protein contains two repeated sequences in its N-terminus and two different repeats in the middle domain. A deletion derivative lacking the C-terminal 187 amino acids remained tightly bound to the cell walls, but no catalytic activity was observed when expressed in *B. subtilis*. By making deletions from the N-terminus it was shown that nearly two-thirds of the protein could be removed without complete loss of cell wall-hydrolyzing activity in *E. coli*, although loss of more than one repeat drastically reduced lytic activity.

[0022] The N-terminal domain of the major autolysin LytA of *Streptococcus pneumoniae* provides the *N*-acetyl-muramyl-L-alanine amidase catalytic function, whereas the C-terminal domain, which contains six repeated sequences, determines the specificity of binding to the cell wall (for review: see reference 18). The protein lacks a signal sequence and requires choline-containing teichoic acids to fully degrade pneumococcal cell walls. Furthermore, it was shown that at least four of the six repeats were needed for efficient recognition of the choline residues of pneumococcal cell walls and the retention of appreciable hydrolytic activity (7).

[0023] LytA, pneumococcal phage lysins as well as clostridial and lactococcal cell wall hydrolases have been used for the construction of active proteins, such that the activity domain and cell wall recognition domains were exchanged. The N-terminal half of the lactococcal phage enzyme was fused to the C-terminal domain of LytA (28). The chimeric enzyme exhibited a glycosidase activity capable of hydrolyzing choline-containing cell walls of *S. pneumoniae*. This result showed that the lactococcal phage lysin consisted of at least two domains with a glucosidase activity contained in its N-terminus and two repeats similar to those in AcmA in the C-terminus (6). A tripartite pneumococcal peptidoglycan hydrolase has been constructed by fusing the N-terminal catalytic domain of the phage CPL1 lysozyme to HBL3, a protein with an amidase activity and a choline-binding domain (27). The three domains acquired the proper conformation as the fusion protein behaved as an amidase, a lysozyme and as a choline-dependent enzyme.

[0024] Also from nature an enzyme is known having two separate functional activity domains: the autolysin gene from *Staphylococcus aureus* encodes a protein that contains an amidase and an endo- β -*N*-acetylglucosaminidase domain separated by three highly similar repeats (20). This protein is processed posttranslationally into the two constituting activity domains.

[0025] The aim of the present study was to investigate the modular structure of AcmA. This was done by consecutively deleting the C-terminal repeats and by fusing the repeats to heterologous proteins. On the basis of cell fractionation and binding studies involving whole cells, it is concluded that the C-terminal repeats in AcmA bind the autolytic enzyme to the cell wall of *L. lactis*.

Materials and Methods

Bacterial strains, plasmids, and growth conditions.

- 5 [0026] The strains and plasmids used in this study are listed in Table 1. *Lactococcus lactis* was grown at 30°C in two-fold diluted M17 broth (Difco Laboratories, Detroit, Mich.) containing 0.5% glucose and 0.95% β -glycerophosphate (Sigma Chemical Co., St. Louis, Mo.) as standing cultures (1/2 M17). Agar plates of the same medium contained 1.5% agar. Five μ g/ml of erythromycin (Boehringer GmbH, Mannheim, Germany) was added when needed. *Escherichia coli* was grown at 37°C with vigorous agitation in TY medium (Difco), or on TY medium solidified with 1.5% agar. When required, the media contained 100 μ g of ampicillin (Sigma), 100 μ g erythromycin or 50 μ g kanamycin (both from Boehringer) per ml. Isopropyl- β -D-thiogalactopyranoside (IPTG) and 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (X-gal) (both from Sigma) were used at concentrations of 1 mM and 0.002%, respectively.

General DNA techniques and transformation.

- 15 [0027] Molecular cloning techniques were performed essentially as described by Sambrook *et al.* (25). Restriction enzymes, Klenow enzyme and T4 DNA ligase were obtained from Boehringer and were used according to the instructions of the supplier. Deoxynucleotides were obtained from Pharmacia (Pharmacia Biotech, Uppsala, Sweden). All chemicals used were of analytical grade and were from Merck (Darmstadt, Germany) or BDH (Poole, United Kingdom). Electrotransformation of *E. coli* and *L. lactis* was performed by using a gene pulser (Bio-Rad Laboratories, Richmond, Calif.), as described by Zabarovsky and Winberg (37) and Leenhouts and Venema (16), respectively. Plasmid DNA was isolated using the QIAGEN plasmid DNA isolation kit (QIAGEN GmbH, Hilden, Germany) or by CsCl-ethidiumbromide density gradient centrifugation and DNA fragments were isolated from agarose gels using the QIAGEN gel extraction kit and protocols from QIAGEN.

Primer synthesis, PCR and DNA Sequencing.

- 25 [0028] Synthetic oligodeoxynucleotides were synthesized with an Applied Biosystems 392 DNA/RNA synthesizer (Applied Biosystems Inc., Foster City, Calif.). The sequences of the oligonucleotides used are listed in Table 2.
- 30 [0029] Polymerase chain reactions (PCR) were performed in a Bio-Med thermocycler 60 (Bio-Med GmbH, Theres, Germany) using super *Tag* DNA polymerase and the instructions of the manufacturer (HT Biotechnology Ltd., Cambridge, United Kingdom). PCR fragments were purified using the nucleotide removal kit and protocol of QIAGEN.
- [0030] Nucleotide sequences of double-stranded plasmid templates were determined using the dideoxy chain termination method (26) with the T7 sequencing kit and protocol (Pharmacia) or the automated fluorescent DNA sequencer 725 of Vistra Systems (Amersham Life Science Inc., Buckinghamshire, United Kingdom).
- 35 [0031] Nucleotide and amino acid sequences were analyzed with the PC/GENE sequence analysis program (version 6.8. IntelliGenetics, Inc., Geneva, Switzerland). Protein homology searches in the SWISSPROT, PIR, and Genbank (release September 23, 1996) databases were carried out with the BLAST program (1).

Construction of *AcmA* derivatives.

- [0032] A stop codon and *EcoRI* restriction enzyme site were introduced in *acmA* at the end of nucleotide sequences encoding the repeats and at the end of the sequence specifying the active site domain by PCR using the primers REPDEL-1, REPDEL-2, and REPDEL-3 and plasmid pAL01 as a template. Primer ALA-4, annealing within the sequence encoding the signal peptide of *AcmA*, was used in all cases as the upstream primer. All three PCR products were digested with *SacI* and *EcoRI* and cloned into the corresponding sites of pBluescript SK+ leading to pDEL1, pDEL2, and pDEL3. Subsequently, the 1,187-bp *PflmI-EcoRI* fragment of pGKAL1 (5) was replaced by the 513, 282 and 76-bp *PflmI-EcoRI* fragments of the inserts of pDEL1, 2 and 3, respectively. The proper plasmids specifying proteins containing one, two or all three repeats (pGKAL5, 4, and 3, respectively) were obtained in *L. lactis* MG1363 *acmA* Δ 1. pGKAL1 was cut with *SpeI*. The sticky ends were flushed with Klenow enzyme and self-ligation introduced a UAG stop codon after the Ser 339 codon of *acmA*. The resulting plasmid was named pGKAL6.
- 50 [0033] A DNA fragment encoding half of the first repeat until the *SpeI* site in the middle of the second repeat was synthesized by PCR using the primers REP-4 A and B. The *NheI* and *SpeI* sites at the ends of the 250-bp PCR product were cut and the fragment was cloned into the unique *SpeI* site of pGKAL1 resulting in plasmid pGKAL7.

Overexpression and isolation of the *AcmA* active site domain.

- [0034] A DNA fragment encoding the active site domain of *AcmA* was obtained using the primers ACMHIS and REP-

DEL-3 with plasmid pAL01 as a template. The 504-bp PCR fragment was digested with *Bgl*II and *Eco*RI and subcloned into the *Bam*HI and *Eco*RI sites of pET32A (Novagen R&D systems Europe Ltd, Abingdon, United Kingdom). The proper construct, pETAcmA, was obtained in *E. coli* BL21(DE3) (30). Expression of the thioredoxin/AcmA fusion protein was induced in this strain by adding IPTG (to 1 mM final concentration) at an OD₆₀₀ of 0.7. Four hours after induction the cells from 1 ml of culture were collected by centrifugation and the fusion protein was purified over a Talon™ metal affinity resin (Clontech Laboratories Inc., Palo Alto, Calif.) using 8 M ureum-elution buffer and the protocol of the supplier. The eluate (200 µl) was dialyzed against a solution containing 50 mM NaCl and 20 mM Tris (pH 7) after which CaCl₂ was added to a final concentration of 2 mM. One unit of enterokinase (Novagen) was added and the mixture was incubated at room temperature for 20 h. The protein mixture was dialyzed against several changes of demineralized water before SDS-PAGE analysis and cell binding studies.

Construction of β-lactamase and α-amylase fusions to the AcmA repeat domain.

[0035] For the introduction of a unique *Nde*I site at the position of the stop codon of the *E. coli* TEM-β-lactamase, the oligonucleotides BETA-1 and BETA-2 were used in a PCR with plasmid pGBL1 (21) as a template. The 403-bp PCR fragment was cut with *Nde*I and *Pst*I and cloned as a 311-bp fragment into the same sites of pUK21. The resulting plasmid, pUKblac, was digested with *Nde*I, treated with Klenow enzyme and subsequently digested with *Xba*I. The β-lactamase encoding fragment was ligated to an 1,104-bp *Pvu*II-*Xba*I DNA fragment from pAL01 containing the *acmA* part encoding the repeat region of AcmA. The resulting plasmid, pUKblacrep, was digested with *Pst*I and *Dra*I and the 1349-bp fragment was inserted into the *Pst*I-*Sna*BI sites of pGBL1: leading to plasmid pGBLR. After digestion of pGAL9 (21) with *Cl*al and *Hind*III the 1,049-bp fragment encompassing the 3'-end of the *Bacillus licheniformis* α-amylase gene was subcloned into corresponding sites of pUK21. According to the paper of Pérez Martínez *et al.* (21), this fragment should be 1,402-bp, but after restriction enzyme analysis it turned out to be approximately 350-bp smaller. The resulting plasmid was called pUKAL1. A unique *Eco*RV restriction enzyme site was introduced by PCR at the position of the stop codon of the *B. licheniformis* α-amylase gene using the oligonucleotides ALFA-A and ALFA-B with plasmid pGAL9 as a template. After restriction of the 514-bp PCR fragment with *Sal*I and *Eco*RV the 440-bp fragment was cloned into the same sites of pUKAL1 resulting in plasmid pUKAL2. The *Eco*RV and *Xba*I sites of this plasmid were used to clone the 1,104-bp *Pvu*II-*Xba*I fragment of pAL01 encoding the repeats of AcmA. The 1,915-bp *Cl*al-*Hind*III fragment of the resulting plasmid pUKALR was used to replace the corresponding 1,049-bp fragment of pGAL9 (pGALR). All cloning steps described above were performed in *E. coli* NM522. The plasmids pGBL1, pGBLR, pGAL9 and pGALR were used to transform *L. lactis* MG1363 and MG1363*acmAΔ1*.

SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and detection of AcmA and α-amylase activity.

[0036] Two ml of end exponential phase *L. lactis* cultures were subjected to centrifugation. 0.5 ml of the supernatant fractions were dialyzed against several changes of demineralized water, lyophilized, and dissolved in 0.25 ml of denaturation buffer (3). Cell pellets were washed with 2 ml of fresh ½M17 medium and resuspended in 1 ml of denaturation buffer. Cell extracts were prepared as described by van de Guchte *et al.* (32).

[0037] AcmA activity was detected by a zymogram staining technique using SDS-PAA (12.5% or 17.5%) gels containing 0.15% autoclaved, lyophilized *Micrococcus lysodeikticus* ATCC 4698 cells (Sigma) as described before (6). For the analysis of α-amylase activity 1% starch was included into 12.5% PAA gels. After electrophoresis proteins were renatured using the AcmA renaturation solution (3) and the gel was stained with an I₂/KI solution (at final concentrations of 12 and 18 mM, respectively) (33).

[0038] SDS-PAGE was carried out according to Laemmli (15) with the Protean II Minigel System (Bio-Rad) and gels were stained with Coomassie brilliant blue (Bio-Rad). The standard low range and prestained low and high range SDS-PAGE molecular weight markers of Bio-Rad were used as references.

[0039] Fractionation of mid- and end-exponential phase cultures of *L. lactis* was performed according to the protocol of Baankreis (2).

Binding of AcmA and its derivatives to lacticoccal cells.

[0040] The cells of 2 ml of exponential phase cultures of MG1363*acmAΔ1* were gently resuspended in an equal volume of supernatant of similarly grown MG1363*acmAΔ1* carrying either plasmid pGK13, pGKAL1, -3, -4, -5, -6 or -7 and incubated at 30°C for 20 min. Subsequently, the mixtures were centrifuged. The cell pellets were washed with 2 ml of ½M17 and cell extracts were prepared in 1 ml of denaturation buffer as described above, while 0.4 ml of the supernatants were dialyzed against demineralized water, lyophilized and dissolved in 0.2 ml of denaturation buffer.

[0041] To analyze competitive binding between AcmA derivatives containing 1 or 2 repeats, equal volumes of the supernatants of MG1363*acmAΔ1* containing pGKAL3 or pGKAL4 were mixed prior to incubation with the

MG1363*acmAΔ1* cells. The samples were treated for SDS-PAGE as described above.

[0042] Three 500 μl samples of a mid-exponential phase culture of MG1363*acmAΔ1* were centrifuged. From one sample 50 μl of the supernatant were replaced by 50 μl of a solution containing the AcmA active site domain (see above). 100 μl of the supernatant of sample two were replaced by 50 μl demineralized water and 50 μl of the supernatant of a mid-exponential phase culture of MG1363*acmAΔ1* (pGKAL4). Of the third sample 100 μl of the supernatant were replaced by 50 μl of the solution containing the AcmA active site domain and 50 μl of the supernatant of MG1363*acmAΔ1* (pGKAL4). Subsequently the three samples were vortexed to resuspend the cells and incubated for 15 min at 30°C. After centrifugation cell and supernatant fractions were prepared in 500 μl of denaturation buffer for analysis of AcmA activity in SDS-(17.5%) PAGE as described above.

[0043] Binding of the β-lactamase/Acma fusion protein was studied by growing MG1363*acmAΔ1* containing pGK13, pGBL1 or pGBLR until mid-exponential phase. The cells of 1 ml of MG1363*acmAΔ1* (pGK13) culture were resuspended in an equal volume of supernatant of either of the other two cultures. The mixtures were prepared in duplo and one series was incubated at 30°C for 5 min while the other was kept at that temperature for 15 min. Then, cell and supernatant fractions were treated as described for the Acma binding studies, resuspended in denaturation buffer in half of the original volume, and subjected to SDS-(12.5%) PAGE followed by Western blot analysis.

Western blotting and immunodetection.

[0044] Proteins were transferred from SDS-PAA gels to BA85 nitrocellulose membranes (Schleicher and Schuell, Dassel, Germany) as described before (31). α-amylase and β-lactamase antigen was detected with 2000-fold diluted rabbit polyclonal anti-ampicillinase antibodies (5 prime→3 prime, Inc., Boulder, Co.), and alkaline phosphatase-conjugated goat anti-rabbit antibodies (Promega Corporation, Madison, Wis.) using the Western-Light Chemiluminescent detection system and protocol (TROPIX Inc., Bedford, Mass.).

Enzyme assays and optical density measurements.

[0045] Acma activity was visualized on 1/2M17 agar plates containing 0.2% autoclaved lyophilized *M. lysodeikticus* cells as halo's around colonies after overnight growth at 30°C.

[0046] α-amylase activity was detected by spotting 10 μl of an overnight culture onto a 1/2M17 agar plate containing 1% of starch (Sigma). After 18 h of incubation at 30°C halo's were visualized by staining with an iodine solution according to the protocol of Smith *et al.* (29). A similar method was used for the detection of β-lactamase activity (29).

[0047] X-prolyl dipeptidyl aminopeptidase (PepX) was measured using the chromogenic substrate Ala-Pro-p-nitroanilid (BACHEM Feinchemicalien AG, Bubendorf, Switzerland). After 2 min of centrifugation in an eppendorf microfuge 75 μl of a culture supernatant was added to 50 μl substrate (2 mM) and 75 μl Hepes buffer (pH 7). The mixture was pipetted into a microtiter plate well and colour development was monitored in a THERMOMax microtiter plate reader (Molecular Devices Corporation, Menlo Oaks, Ca) at 405 nm during 20 minutes at 37°C. Optical densities were measured in a Novaspec II spectrophotometer (Pharmacia) at 600 nm.

Results

Two of the three repeats in Acma are sufficient for autolysis and cell separation.

[0048] Several mutant Acma derivatives were constructed to investigate the function of the three repeats in the C-terminus of Acma. A stop codon was introduced behind the codon for Thr-287 (pGKAL4) or Ser-363 (pGKAL3) (see Fig. 1). Plasmid pGKAL4-specified Acma (A1) only contains the first (most N-terminal) of the three repeats, while pGKAL3 specifies an Acma variant (A2) carrying the first two repeats. pGKAL5 specifies an Acma derivative lacking repeats (A0) due to the introduction of a stop codon after Ser-218. Acma specified by pGKAL6 contains one and a half repeat (A1.5) due to the presence of a stop codon behind the Ser-339 codon. From pGKAL7 an Acma mutant (A4) is produced which carries an additional (fourth) repeat as the result of duplication of the polypeptide from Ser-263 to Thr-338. All proteins were expressed from the *acmA* promoter in the Acma-negative strain *L. lactis* MG1363*acmAΔ1*. The various deletions of Acma were examined with respect to the following properties: (I) their effect on halo formation on plates containing cell wall fragments of *M. lysodeikticus*, (II) chain length of the cells expressing the mutant Acma's, and sedimentation of the cells in a standing culture, (III) their enzymatic activity, both in the cell and supernatant fraction and (IV) autolysis.

[0049] Halo formation. On a 1/2M17 plate containing cell wall fragments of *M. lysodeikticus* halo's were absent when MG1363*acmAΔ1* carried pGK13 or pGKAL5. All other strains produced a clear halo that differed in size. The halo size was clearly correlated with the number of full length repeats present, although the addition of an extra repeat resulted in a reduced halo size (see Table 3). Apparently, for optimal cell wall lytic activity a full complement of repeats is

required.

[0050] *Cell separation and sedimentation.* The deletion of one and a half, two and all three repeats had a clear effect on the chain length and on sedimentation of the cells after overnight growth (see Table 3). Thus, efficient cell separation requires the presence of at least two repeats in AcmA.

5 [0051] *Enzyme activity.* Cells and supernatants of overnight cultures of all strains were analyzed for AcmA activity by SDS-PAGE. In the cell fractions no activity was detected for A0, not even after one week of renaturation of the protein (Table 3). Of the other derivatives, two major activity bands were present in this fraction. In each case their positions in the gel corresponded to proteins with the calculated molecular weights of the unprocessed and the processed form. (Table 3 and not shown). As shown in Fig. 1, all AcmA derivatives were still active in the supernatant fractions. AcmA 10 produced the characteristic breakdown pattern as determined before (Fig. 1, lanes 1 and 3; (6)). All AcmA derivatives except A0 and A1 also showed a distinct and highly reproducible degradation pattern. A4 showed 2 additional breakdown products after prolonged renaturation (results not shown). These data indicate that removal of the repeats does not destroy AcmA activity and suggests that one repeat is sufficient to keep the enzyme cell-associated.

[0052] *Autolysis.* To analyze the effect of the repeats on autolysis during stationary phase, overnight cultures of all 15 strains were diluted hundred-fold and incubated at 30°C for 6 days and the decrease of optical density (OD₆₀₀) was followed. All cultures exhibited similar growth rates, reached the same maximal optical densities and did not lyse during the exponential phase of growth. After approximately 60 h of incubation maximal reduction in OD₆₀₀ was reached in all cases. The results are presented in Table 3 and show that the reduction in OD₆₀₀ is correlated with the reduction of the number of AcmA repeats. To investigate whether the decrease in OD₆₀₀ really reflected autolysis, the activity of the 20 intracellular enzyme PepX was measured. After 60 h of incubation, PepX activity in the culture medium was also maximal in all samples, decreasing in all cases upon further incubation. Hardly any PepX activity was detected in the supernatant of the *acmAΔ1* mutant and in cultures producing A0, A1 or A1.5. In contrast, a considerable quantity of PepX had released into the supernatant of cultures producing A2 and A3. Thus two repeats in AcmA are sufficient for autolysis of *L. lactis*. A2 or A4 production led to reduced lysis of the producer cells. Taken together, these results indicate 25 that the repeats in AcmA function in efficient autolysis and are required for cell separation.

The active site domain of AcmA resides in the N-terminal part.

[0053] To examine whether the active site is located in the N-terminal domain of AcmA, a DNA fragment starting at 30 codon 58 until codon 218 of *acmA* was synthesized by PCR and fused to the thioredoxin gene in plasmid pET32A. The fusion protein comprises 326 amino acids. A protein with the expected molecular mass (35 kDa) was isolated from a culture of *E. coli* BL21(DE3) (pETAcmA) (Fig. 2, lane A2). By cleavage with enterokinase, the protein was split into a thioredoxin part of 17 kDa and an AcmA domain (nA) of 18 kDa (Fig. 2, lane A1). The zymogram (Fig. 2.B) shows that 35 the fusion protein did not have appreciable cell wall hydrolytic activity, while the released domain of AcmA was active (Fig. 2, lanes B1 and B2), indicating that the active site domain was in the N-terminal part of AcmA.

Fusion of the repeats of AcmA to α -amylase and β -lactamase yields active enzymes.

[0054] The three C-terminal repeats of AcmA (cA) were fused C-terminally to *B. licheniformis* α -amylase and *E. coli* 40 TEM β -lactamase as described in Material and Methods and shown in Fig. 3. The hybrid proteins were fused to the lactococcal signal sequences AL9 and BL1, respectively (21). Both fusion proteins were active in plate assays, as is only shown for the β -lactamase/AcmA fusion protein (β cA) (Fig. 4). The halo's around colonies producing the fusion proteins were smaller than those produced by the wild-type enzymes, which could either be caused by reduced intrinsic enzyme activities due to the presence of repeats or by increased susceptibility to proteolytic degradation. However, the smaller 45 halo's produced by the chimeric proteins might also be caused by hampered diffusion due to cell wall binding (see below). The activities of α -amylase and the α cA fusion protein were also detected in a renaturing SDS-(12.5%)PAA gel containing 1% starch. The primary translation product of the α -amylase gene is a protein of 522 amino acid residues which contains a signal sequence of 37 amino acids (21). It is secreted as a 55-kDa protein. α cA consists of 741 amino acids and, if processed and secreted, would give rise to a 78-kDa protein. Cell and supernatant fractions of *L. lactis* 50 MG1363 and MG1363*acmAΔ1* carrying pGAL9 or pGALR were analyzed after overnight growth of the strains. The results are presented in Fig. 5 and show that the clearing bands are present at the position expected for both mature proteins. Apparently, α cA is active. Clearly, smaller products are present in the supernatants of the cells producing the fusion protein, the smallest being approximately of the size of wild-type mature α -amylase (Fig. 5 and not shown). The β -lactamase fusion protein is predominantly present in the cell wall.

55 [0055] To examine whether the presence of the C-terminal domain of AcmA resulted in binding of β cA to the cell wall, mid-exponential phase cultures of *L. lactis* MG1363*acmAΔ1* containing pGBL1, encoding β -lactamase, or pGBLR, specifying β cA, were fractionated and subjected to Western blot analysis (Fig. 6). From pGBL1, β -lactamase is expressed as a protein of 322 amino acids containing a signal sequence of 47 amino acids. The secreted protein is 30

kDa. β ca consists of 540 amino acids and is secreted as a protein with a molecular mass of 52 kDa. Figure 6 shows that most of the wild-type β -lactamase is present in the culture supernatant and none in the cytoplasm. Slightly larger bands, likely representing the unprocessed form, are found in the membrane fractions of this strain. In contrast, β ca is predominantly retained in the cell wall fraction, although a considerable amount resides in the cytoplasm, strongly suggesting that the AcmA repeats anchored the hybrid enzyme to the cell wall. The smaller band present in both cytoplasmic fractions is caused by cross hybridization of the antibodies to an unspecified lactococcal protein (unpublished observation). In the supernatant fraction of cells producing β ca, only little full length protein was observed. Several distinct smaller products are present in this fraction which were also detectable in very low amounts in the cell wall fraction after prolonged exposure of the film (not shown) but were absent from the other fractions.

The C-terminal repeats in AcmA are required for cell wall binding.

[0056] Although the results presented in the previous section strongly suggests that the C-terminal repeats are required for the retention of protein in the cell wall, definite proof was obtained by mixing the supernatant fractions of end-exponential phase cultures containing AcmA, or one of its deletion derivatives (see Fig. 1), with the cells from an equal volume of a culture of MG1363acmA Δ 1 (pGK13). After incubation, cell and supernatant fractions were examined for the presence of AcmA. Except for A0, all proteins were capable of binding to the MG1363acmA Δ 1 cells (Table 3). Also, all degradation products of AcmA and its derivatives were capable of binding. The finding that A0 was unable to bind was corroborated by adding the mixture of enterokinase-released nA and thioredoxin to supernatant containing A1. When incubated with AcmA-minus cells, only A1 bound to the lactococcal cells (Fig. 7) as only this protein was detectable in the cell fraction. nA was only detected in the supernatant. This was also the case when the experiment was repeated with nA alone (not shown).

Binding of AcmA or β ca to lactococcal cells at different pHs.

[0057] The supernatant fraction of a mid-exponential phase *L. lactis* MG1363acmA Δ 1 culture was replaced by the supernatant of a mid-exponential phase *L. lactis* MG1363 culture. This mixture was incubated at 30°C for 5 min. Thereafter the supernatant was removed by centrifugation and the cell pellet was washed with M17. The cell pellets were dissolved in M17 with pHs ranging from 2 to 10 and incubated at 30°C for 30 min. The cell and supernatant fractions were separated and treated as described before and analysed for the presence of AcmA activity. A similar experiment was executed with mid-exponential phase *L. lactis* MG1363acmA Δ 1 cells with the supernatant of an *L. lactis* MG1363acmA Δ 1(pGBLR) culture. The presence of β ca was analysed by western blotting and immunodetection as described.

[0058] At all different pHs, AcmA and β ca was found to be bound to the lactococcal cells. The binding of both AcmA and β ca was better at low pH as judged from the activity in a zymogram and the visual presence of the amount of β ca fusion protein in the cell extracts after immunodetection.

Proteolytic breakdown of AcmA by Pronase and Trypsin.

[0059] The supernatant fraction of a mid-exponential phase MG1363acmA Δ 1 culture was replaced by the supernatant of a mid-exponential phase MG1363 culture. This mixture was incubated at 30°C for 15 min. Thereafter the supernatant and the cell fractions were separated and the cell pellet was dissolved in an identical volume of M17. To both fractions Pronase and Trypsin (1 mg/ml) dissolved in 10 mM NaPi buffer (pH=7) was added to an end concentration of (10 μ g/ml) and the mixtures were incubated at 30°C. Samples were taken after 5 and 30 min and 2 h of incubation. The cell and supernatant fractions of each sample were separated and prepared for zymographic analysis as described above.

[0060] A complete hydrolysis of AcmA by pronase was observed in the supernatant fraction after 2 h of incubation while activity was still present in the cell extract at this time point. The hydrolysis of AcmA by trypsin was slower and activity was still present in the supernatant after 2 h of incubation. In time the portion of activity present in the cell extracts was always higher than that observed in the supernatant. These results indicate that the AcmA protein is protected when it is bound to the cell.

Binding of AcmA to different types of bacterial cells.

[0061] The strains *Bacillus subtilis* DB104, *Lactobacillus plantarum* 80, *Streptococcus faecalis* JH2-2, *Streptococcus thermophilus* ATCC 19258, *Listeria P*, *Lactobacillus buchneri* L4, *Clostridium beijerinckii* CNRZ 530 and *Escherichia coli* NM522 were grown overnight in GM17. Two fractions of each overnight culture were centrifuged and the supernatants were replaced by the supernatant of an overnight-culture of *L. lactis* MG1363acmA Δ 1(pGKAL1) or MG1363acmA Δ 1(pGK13). The mixtures were incubated at 30°C for 15 min. Subsequently the cell and supernatant fractions were separated and the cells were washed once with M17 and were prepared for SDS-PAGE as described before and analysed for AcmA activity.

[0062] In all cell extracts AcmA activity was present while such an activity was absent in extracts of cells which had

been incubated with the supernatant of MG1363*acmA*Δ1(pGK13) which lacks the presence of AcmA.

[0063] To investigate the effect of repeat number on binding, equal volumes of the supernatants of cultures of MG1363*acmA*Δ1 (pGKAL3, encoding A2) and MG1363*acmA*Δ1 (pGKAL4, specifying A1) were mixed. The undiluted and a 10-fold diluted mixture were incubated with the AcmA-free cells. Analysis of zymograms of serial dilutions showed that the two activities were equally distributed over the cell and supernatant fractions, indicating that both proteins bind equally well (results not shown).

[0064] To examine whether the C-terminal repeat sequences of AcmA had the capacity to bind a heterologous, extra-cellular enzyme to lactococcal cells, binding of βcA was assessed by incubation of AcmA-minus *L. lactis* cells with culture supernatants containing either secreted wild-type β-lactamase or βcA. As Fig. 8 shows, wild-type β-lactamase was exclusively present in the supernatant fraction, whereas βcA fractionated with the lactococcal cells and, thus, had bound to these cells.

Discussion

[0065] The results presented in this work indicate that the mature form of the *N*-acetylmuramidase AcmA of *L. lactis* consists of two separate domains. The overproduced and purified N-terminus, from amino acid residue 58 to 218 in the pre-protein, is active on *M. lysodeikticus* cell walls and, thus, contains the active site of the enzyme. This is in agreement with the finding that the repeat-less AcmA mutant A0 can still hydrolyze *M. lysodeikticus* cell walls, albeit with severely reduced efficiency. Prolonged renaturation was needed to detect the activity of the enzyme *in vitro* while colonies producing the protein did not form a halo. Enzymes A1 and A2 had *in vitro* activities which were nearly the same as that of the wild-type protein, although in the plate assay A1 produced a smaller halo than A2 which, in turn, was smaller than the wild-type halo. A strain producing A1 grew in longer chains than cells expressing A2 and, in contrast to A2 producing cells, sedimented and did not autolyse. Taken together these results indicate that, although the N-terminus of AcmA contains the active site, the presence of at least one complete repeat is needed for the enzyme to retain appreciable activity. Second, only cultures producing AcmA's containing two or more full length repeats are subject to autolysis and produce wild-type chain lengths. It is tempting to speculate that this apparent increase in catalytic efficiency of AcmA is caused by the repeat domain by allowing the enzyme to bind to its substrate, the peptidoglycan of the cell wall. As was postulated by Knowles *et al.* (12) for the cellulase binding domains in cellobiohydrolases, such binding would increase the local concentration of the enzyme. The repeats could be involved in binding alone or could be important for proper positioning of the catalytic domain towards its substrate. The increase in AcmA activity with an increasing number of repeats to up to 3 in the wild-type enzyme, suggests an evolutionary process of repeat amplification to reach an optimum for proper enzyme functioning. The binding of A1, A1.5 and A4 was comparable with that of wild-type AcmA but these enzyme varieties caused no or only little autolysis. These observations seem to support the idea that 3 repeats are optimal for proper functioning of AcmA. The presence of 5 and 6 repeats in the very similar enzymes of *E. faecalis* and *E. hirae*, respectively, may reflect slight differences in cell wall structure and/or the catalytic domain, requiring the recruitment by these autolysins of extra repeats for optimal enzyme activity.

[0066] The hypothesis that the C-terminal domain of AcmA is involved in cell binding (6) was corroborated in this study. First of all we show that AcmA is indeed capable of cell binding. AcmA and its derivatives A1, A1.5, A2 and A4 all bound to cells of *L. lactis* when added from the outside. To prove that it was the C-terminus of AcmA that facilitated binding and not some intrinsic cell wall binding capacity of the N-terminal domain, the repeat domain was fused to two heterologous proteins which do not normally associate with the cell wall. The smaller halo's produced by αcA and βcA compared to the wild-type proteins and the presence of most of βcA in the cell wall fraction are indicative of cell binding of the fusion proteins via the AcmA-specific repeats.

[0067] The βcA binding studies clearly show that it is the AcmA repeat domain that specifies cell wall binding capacity: whereas wild-type β-lactamase (and, for that matter, repeat-less AcmA) did not bind to lactococcal cells, βcA did bind to these cells when added from the outside. The results obtained with A1 in the binding assay show that only one repeat is sufficient to allow efficient binding of AcmA. In a separate study (5) we showed that AcmA can operate intercellularly: AcmA-free lactococcal cells can be lysed when grown together with cells producing AcmA. Combining this observation with the results presented above allows to conclude that AcmA does not only bind when confronting a cell from the outside but, indeed, is capable of hydrolyzing the cell wall with concomitant lysis of the cell.

[0068] AcmA-like repeats were found to be present at different locations in more than 30 proteins after a comparison of the amino acid sequences of the repeats in AcmA with the protein sequences of the Genbank database (release 23). Not all of these proteins with repeats varying from one to six are cell wall hydrolases. Alignment of the amino acid sequences of all the repeats yielded a consensus sequence similar to that postulated by Birkeland and Hourdou *et al.* (4, 9). Interestingly, if a limited number of modifications are allowed in the consensus repeat, the repeat is also present 12 and 4 times, respectively, in two proteins of *Caenorhabditis elegans*, which both show homology with endochitinases (Gene accession numbers U64836 and U70858) (36). Possibly, these repeats anchor these enzymes to fungi ingested by this organism. The presence of similar repeats in proteins of different bacterial species strongly suggests that they

recognize and bind to a general unit of the peptidoglycan. An interesting goal for the future will be to elucidate the unit to which they bind and the nature of the binding.

[0069] As has been reported earlier for intact AcmA (5), and, as we show here for its C-terminal deletion derivatives, the enzyme is subject to proteolytic degradation. None of the degradation products were present in cell extracts of whole cells indicating that they are not formed inside the cell (data not shown). The degradation pattern of each AcmA derivative is specific and very reproducible. Based on the sizes of the degradation products, a number of the proteolytic cleavage sites probably resides in the intervening sequences. One such site (1 in Fig. 1) is present between repeat 1 and 2. Cleavage at this position would result in an active protein of approximately 28 kDa, which is indeed seen in the supernatants of all strains producing AcmA with 1.5 or more repeats. A second cleavage site is probably located between the second and third repeat (2 in Fig. 1). Cleavage at this site is either rather infrequent, or the resulting degradation product is not very active, which, in both cases, would lead to the faint bands of activity observed in lanes 1 and 3 of the zymogram presented in Fig. 1. The presence of cleavage sites in between the AcmA repeats is further suggested by the presence of specific degradation products observed in α cA and β cA; their sizes are in accord with the location of the cleavage sites postulated in AcmA. In addition, as also bands of the size of the wild-type α -amylase and β -lactamase are observed, an additional cleavage site seems to be present around the fusion point of these enzymes and the cell wall binding domain of AcmA.

[0070] All degradation products of AcmA and those of the two fusion proteins are mainly present in the supernatant and to some extent in the cell wall fraction, but not in the cells. As none of the *L. lactis* strains used produced the cell wall-anchored proteinase PrtP, this enzyme can not be held responsible for the specific degradation of AcmA or the fusion proteins. Apparently, an extracellular proteinase exists in *L. lactis* that is capable of removing the repeats, which may represent a mechanism for the regulation of AcmA activity.

TABLE 1. Bacterial strains and plasmids used in this study.

Strain or plasmid	Relevant phenotype(s) or genotype(s)	Source of reference
Strains		
<i>L. lactis</i> subsp. <i>cremoris</i>		
MG1363	Plasmid-free strain	(8)
MG1363acmA41	Derivative of MG1363 carrying a 701-bp <i>SacI</i> - <i>SpeI</i> deletion in <i>acmA</i>	(6)
<i>E. coli</i>		
NM522	<i>supE</i> <i>thi</i> Δ (<i>lac</i> - <i>proAB</i>) Δ <i>hds5</i> (r_k^+ , m_k^+) [<i>F'</i> <i>proAB lacI^qZM15</i>]	Stratagene (30)
BL21 (DE3)	<i>ompT</i> r_s^+ m_g^- <i>int</i> ; bacteriophage DE3 lysogen carrying the T7 RNA polymerase gene controlled by the <i>lacUV5</i> promoter	
Plasmids		
pET32A	<i>Ap^r</i> , vector for high level expression of thioredoxin fusion proteins	Novagen
pUK21	<i>Km^r</i> , general cloning vector	(35)
pBluescript SK+	<i>Ap^r</i> , general cloning vector	Stratagene
pAL01	<i>Ap^r</i> , pUC19 carrying a 4,137-bp lactococcal chromosomal DNA insert with <i>acmA</i> gene	(6)
pDEL1	<i>Ap^r</i> , pBluescript SK+ with 785-bp <i>SacI</i> - <i>EcoRI</i> fragment of <i>acmA</i> obtained by PCR with primers ALA-4 and REPDEL-1	This work
pDEL2	<i>Ap^r</i> , pBluescript SK+ with 554-bp <i>SacI</i> - <i>EcoRI</i> fragment of <i>acmA</i>	This work

pDEL3	obtained by PCR with primers ALA-4 and REPDEL-2 Ap ^r , pBluescript SK+ with 348-bp SacI-EcoRI fragment of acmA	This work
pGKAL1	obtained by PCR with primers AIA-4 and REPDEL-3 Em ^r , Cm ^r , pGK13 containing acmA under control of its own promoter on a 1,942-bp SspI-BamHI insert	(5)
pGKAL3	Em ^r , Cm ^r , pGKAL1 derivative expressing A2	This work
pGKAL4	Em ^r , Cm ^r , pGKAL1 derivative expressing A1	This work
pGKAL5	Em ^r , Cm ^r , pGKAL1 derivative expressing A0	This work
pGKAL6	Em ^r , Cm ^r , pGKAL1 derivative expressing A1.5	This work
pGKAL7	Em ^r , Cm ^r , pGKAL1 derivative expressing A4	This work
pETACMA	Ap ^r , pET32A expressing active site domain of AcMA from residues 58 to 218 fused to thioredoxin	This work
pGBL1	Em ^r , pWV01 derivative expressing <i>E. coli</i> TEM- β -lactamase fused to export element BL1 of <i>L. lactis</i>	(21)
pGAL9	Em ^r , pWV01 derivative expressing <i>B. licheniformis</i> α -amylase fused to export element AL9 of <i>L. lactis</i>	(21)
pUKAL1	Km ^r , pUK21 with \pm 1,050-bp ClaI-HindIII fragment of pGAL9	This work
pUKAL2	Km ^r , pUKAL1 in which the \pm 650-bp SalI-EcoRV fragment is replaced by the 440-bp SalI-EcoRV fragment of the PCR fragment obtained with primers ALFA-A and -B	This work
pUKALR	Km ^r , pUKAL2 with 1,104-bp PvuII-XbaI fragment of pAL01 in EcoRV and XbaI sites	This work

5
10
15
20
25
30
35
40
45
50
55

pUKblac	Km ^r , pUK21 with 311-bp PstI-NdeI PCR fragment obtained with primers BETA-1 and -2	This work
pUKblacR	Km ^r , pUKblac carrying 1,104-bp PvuII-XbaI fragment of pAL01 in NdeI and XbaI sites	This work
pGBLR	Em ^r , pGBL1 expressing the β -lactamase/AcmA fusion protein	This work
pGALR	Em ^r , pGAL9 expressing the α -amylase/AcmA fusion protein	This work

Table 2

Oligonucleotides used in this study ^{a)}		
Name	Nucleotide sequence (5'→3')	R/E site
REPDEL-1	CGCGAATTCAGATTATGAAACAATAAG	<i>EcoRI</i>
REPDEL-2	CGCGAATTCCTTAGTACAAGTTTTTG	<i>EcoRI</i>
REPDEL-3	CGCGAATTCCTTATGAAGAAGCTCCGTC	<i>EcoRI</i>
ALA-4	CTTCAACAGACAAGTCC	
REP-4A	AGCAATACTAGTTTTATA	<i>SpeI</i>
REP-4B	CGCGAATTCGCTAGCGTCGCTCAAATTCAAAGTGCG	<i>NheI</i>
ACMHIS	AGGAGATCTGCGACTAACTCATCAGAGG	<i>BglII</i>
BETA-1	GGATCATGTAACTCGCC	
BETA-2	GGAATTCATATGCTTAATCAGTGAGG	<i>NdeI</i>
ALFA-A	GCA ^T CCGTTGAAAGCGG	
ALFA-B	GAATTCGATATCTTTGAACATAAATTG	
ALA-14	GATAAATGATTCCAAGC	<i>EcoRV</i>
ALA-22	CTCAAATTCAAAGTGCG	

a) The indicated restriction enzyme (R/E) sites are underlined while stopcodons are shown in italic.

Table 3. Properties of *L. lactis* expressing AcMA derivatives.

1	MG pGK13	A3	32.6	16.9	A	3.1	-	+	+	+
2	Δ1 pGK13	-	15.2	0.3	C	0	+	-	-	-
3	Δ1 pGKAL1	A3	36.7	19.8	A	5.0	-	+	+	+
4	Δ1 pGKAL3	A2	29.3	13.3	A	4.6	-	+	+	+
5	Δ1 pGKAL4	A1	18.8	0.4	B	3.9	+	+	+	+
6	Δ1 pGKAL5	A0	15.6	0.3	C	0	+	+	-	-
7	Δ1 pGKAL6	A1.5	18.6	1.6	B	2.2	±	+	+	+
8	Δ1 pGKAL7	A4	21.1	4.9	A	4.0	-	+	+	+

- a) The number corresponds to the AcmA derivative produced, as schematized in Fig. 1.
- b) MG: *L. lactis* MG1363, Δ1: *L. lactis* MG1363acmΔ41.
- c) -: no AcmA produced; Ax: AcmA with x repeats.
- d) The OD_{600} reduction was calculated using the following formula: $[(OD_{max} - OD_{60\text{ heures}}) / OD_{max}] * 100\%$.
- e) Activity is in arbitrary units measured as the increase of absorption at 405 nm in time.
- f) End exponential phase $\frac{1}{2}$ M17 cultures were subjected to light microscopic analysis.
 - A: mainly single cells and some chains up to 5 cells
 - B: some single cells but mainly chains longer than 5 cells
 - C: no single cells, only very long chains
- g) The sizes of the halo's were measured in millimeters from the border of the colony after 45 h of incubation at 30°C.
- h) Analyzed by visual inspection of standing $\frac{1}{2}$ M17 cultures after overnight growth in test tubes.
- i) Judged from zymograms of samples from end-exponential phase $\frac{1}{2}$ M17 cultures; sup: supernatant fraction, ce: cell-extract.
- j) Binding of AcmA derivatives in supernatants of end-exponential phase $\frac{1}{2}$ M17 cultures to end-exponential phase cells of *L. lactis* MG1363acmΔ41 after 20 min of incubation at 30°C (see text for details).

FIG. 1. Analysis of AcmA activity in supernatant fractions of end-exponential-phase cultures of MG1363 containing

pGK13 (1) and MG1363*acmA*Δ1 containing either pGK13, not encoding AcmA (2), pGKAL1, encoding enzyme A3 (3), pGKAL3, encoding enzyme A2 (4), pGKAL4, encoding enzyme A1 (5), pGHAL5, encoding enzyme A0 (6), pGKAL6, encoding enzyme A1.5 (7), or pGKAL7, encoding enzyme A4 (8) in a renaturing SDS-(12.5%)PAA gel containing 0.15% *M. lysodeikticus* autoclaved cells. Molecular masses (in kilodaltons, kDa) of standard proteins (lane M) are shown in the left margin. Below the gel the lower part of lanes 5, 6 and 7 of the same gel is shown after one week of renaturation. The right half of the figure gives a schematic representation of the various AcmA derivatives. SS (black), signal sequence; Rx (dark grey), repeats; light grey, Thr, Ser and Asn-rich intervening sequences (6); arrows, artificially duplicated region in the AcmA derivative containing four repeats. The active site domain is shown in white. MW, expected molecular sizes in kDa of the secreted forms of the AcmA derivatives. The numbers of the AcmA derivatives correspond with the lane numbers of the gel. Numbered arrowheads indicate the putative location of proteolytic cleavage sites.

FIG. 2. Purification of the AcmA active site domain (nA). (A) SDS-(12.5%) PAGE of cell extract of 10 μl of *E. coli* BL21(DE3) (pETA*acmA*) (lane 3) induced for 4 h with IPTG. Lane 2, 10 μl of purified fusion protein isolated from 25 μl of induced *E. coli* culture and lane 1, 10 μl of the enterokinase cleft protein. (B) Renaturing SDS-(12.5%) PAGE with 0.15% *M. lysodeikticus* autoclaved cells using the same amount of the samples 1 and 2 shown in part A. Molecular masses (in kilodaltons) of standard proteins are shown on the left of the gel. Before loading the samples were mixed with an equal volume of 2x sample buffer (15).

FIG. 3. Schematic representation of plasmids pGBLR and pGALR carrying, respectively, C-terminal fusions of the repeats of AcmA to β-lactamase and α-amylase. α-amy, α-amylase gene of *B. licheniformis*; β-lac, β-lactamase gene of *E. coli*; *acmA*, 3-prime end of the *N*-acetylmuramidase gene of *L. lactis* MG1363 encoding the three repeats; *EmR* and *ChlR*, erythromycin and chloramphenicol resistance genes; AL9 and BL1, protein secretion signals from *L. lactis* MG1363 (21); *repA* and ORI, gene for the replication protein and origin of replication of the lactococcal plasmid pWV01, respectively; Psp02, *B. subtilis* phage Spo2 promoter. Black boxes indicate the PCR fragments used for the introduction of the restriction enzyme sites *EcoRV* and *NdeI* at the position of the stopcodons of the α-amylase and β-lactamase genes, respectively. The open box indicates the part which has been subcloned into pUK21 for construction work. The grey boxes show the fragment of pAL01 used to fuse the 3'-end of *acmA* to the α-amylase and β-lactamase genes. Only relevant restriction enzyme sites are shown.

FIG. 4. β-lactamase activity in *L. lactis*. Activity of wild-type β-lactamase and its AcmA fusion derivative (βcA) produced by cells of *L. lactis* MG1363 and MG1363*acmA*Δ1 containing pGK13, pGBL1 or pGBLR. The 1/2M17 agar plate was stained with iodine after overnight growth of the colonies according to the protocol of Smith *et al.* (29).

FIG. 5. α-amylase activity in the supernatant of *L. lactis*. Activity of wild-type α-amylase (α) and the αcA fusion protein in an SDS-(12.5%)PAA gel containing 1% starch. The proteins were renatured by washing the gel with Triton X-100 and subsequently stained with iodine (33). The equivalent of 40 μl of supernatant of 1/2M17 cultures of *L. lactis* MG1363 (M) and MG1363*acmA*Δ1 (Δ) containing pGAL9 or pGALR were loaded onto the gel. Molecular masses (in kDa) of standard proteins are shown in the left margin.

FIG. 6. Localization of β-lactamase in *L. lactis*. Western blot analysis of fractions of MG1363*acmA*Δ1 expressing β-lactamase (from pGBL1) or βcA fusion protein (encoded by pGBLR) using polyclonal antibodies directed against β-lactamase. Amount of samples loaded is equal to 200 μl of culture. Fractions: S, supernatant; CW, cell wall; CY, cytoplasm; MB, membrane-associated; and M, membrane.

FIG. 7. Analysis of the binding of AcmA derivatives nA and A1 by a renaturing SDS-(17.5%) PAGE with 0.15% *M. lysodeikticus* autoclaved cells. Cell (C) and supernatant (S) fractions of MG1363*acmA*Δ1 cells incubated with nA and A1 from the culture supernatant of MG1363*acmA*Δ1 containing pGKAL4. 60 μl of the samples were loaded. Molecular masses (in kDa) of standard proteins are shown in the left margin.

FIG. 8. Binding of the βcA fusion protein to *L. lactis*. The figure shows a Western blot using polyclonal antibodies against β-lactamase. Cell extracts (lanes 1, 3, 5) and supernatants (lanes 2, 4, 6) of mid-exponential phase MG1363*acmA*Δ1 (pGK13) cells incubated for 5 minutes with supernatants of MG1363*acmA*Δ1 containing pGK13 (lanes 1, 2), pGBLR (lanes 3, 4) or pGBL1 (lanes 5, 6), respectively. The positions of wild-type β-lactamase (β) and the βcA fusion protein are indicated on the right. Molecular masses (in kDa) of standard proteins are shown in the left margin. Twenty μl of samples were loaded onto an 12.5% PAA gel.

FIG 9. Schematic representation of the AcmA protein. SS (black), signal sequence; R (dark grey), repeats; shaded

regions, intervening sequences. The active site domain is shown in white.

FIG. 10. Amino acid sequence alignment of the repeats of AcmA in *L. lactis* plus consensus sequence.

Fig 11. Amino acid sequence alignment of repeats in various species

- a) Proteins listed were obtained by a homology search in the SWISSPROT, PIR, and Genbank databases with the repeats of AcmA using the BLAST program (1).
- b) * ; genes encoding cell wall hydrolases.
- # ; proteins containing repeats that are longer than average
- c) The number of aa residues between the repeats are given between brackets.
- d) Number of aa of the primary translation product.
- e) Genbank accession number.

References

[0071]

1. Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403-410.
2. Baankrels, R. 1992. The role of lactococcal peptidases in cheese ripening. PhD Thesis. University of Amsterdam, the Netherlands.
3. B  lveau, C., C. Potvin, J. Trudel, A. Asselin, and G. Bellemare. 1991. Cloning, sequencing, and expression in *Escherichia coli* of a *Streptococcus faecalis* autolysin. *J. Bacteriol.* **173**:5619-5623.
4. Birkeland, N.-K. 1994. Cloning, molecular characterization, and expression of the genes encoding the lytic functions of lactococcal bacteriophage   LC3: a dual lysis system of modular design. *Can. J. Microbiol.* **40**:658-665.
5. Buist, G., H. Karsens, A. Nauta, D. Van Sinderen, G. Venema, and J. Kok. 1997. Autolysis of *Lactococcus lactis* caused by induced overproduction of its major autolysin AcmA. *Appl. Environ. Microbiol.* **63**: 2722-2728.
6. Buist, G., J. Kok, H. J. Leenhouts, M. Dabrowska, G. Venema, and A. J. Haandrikman. 1995. Molecular cloning and nucleotide sequence of the gene encoding the major peptidoglycan hydrolase of *Lactococcus lactis*, a muramidase needed for cell separation. *J. Bacteriol.* **177**:1554-1563.
7. Garc  a, J. L., E. D  az, A. Romero, and P. Garc  a. 1994. Carboxy-terminal deletion analysis of the major pneumococcal autolysin. *J. Bacteriol.* **176**:4066-4072.
8. Gasson, M. J. 1983. Plasmid complements of *Streptococcus lactis* NCDO 712 and other lactic Streptococci after protoplast-induced curing. *J. Bacteriol.* **154**:1-9.
9. Hourdou, M.-L., M. Guinand, M.-J. Vacheron, G. Michel, L. Denoroy, C. Duez, S. Englebert, B. Joris, G. Weber, and J.-M. Ghuysen. 1993. Characterization of the sporulation-related γ -D-glutamyl-(L)-meso-diaminopimelic-acid-hydrolysing peptidase I of *Bacillus sphaericus* NCTC 9602 as a member of the metallo(zinc) carboxypeptidase A family (Modular design of the protein). *Biochem. J.* **292**:563-570.
10. Joris, B., S. Englebert, C. P. Chu, R. Kariyama, L. Daneo-Moore, G. D. Shockman, and J. M. Ghuysen. 1992. Modular design of the *Enterococcus hirae* muramidase-2 and *Streptococcus faecalis* autolysin. *FEMS Microbiol. Lett.* **91**:257-264.
11. Kariyama, R. and G. D. Shockman. 1992. Extracellular and cellular distribution of muramidase-2 and muramidase-1 of *Enterococcus hirae* ATCC 9790. *J. Bacteriol.* **174**:3236-3241.
12. Knowles, J., P. Lehtovaara, and T. Teeri. 1987. Cellulase families and their genes. *Tibtech.* **5**:255-261.
13. Kuroda, A. and J. Sekiguchi. 1990. Cloning, sequencing and genetic mapping of a *Bacillus subtilis* cell wall hydrolase gene. *J. Gen. Microbiol.* **136**:2209-2216.
14. Kuroda, A., Y. Sugimoto, T. Funahashi, and J. Sekiguchi. 1992. Genetic structure, isolation and characterization of a *Bacillus licheniformis* cell wall hydrolase. *Mol. Gen. Genet.* **234**:129-137.
15. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**:680-685.
16. Leenhouts, K. and G. Venema. 1992. Molecular cloning and expression in *Lactococcus*. *Med. Fac. Landbouww. Univ. Gent.* **57**:2031-2043.
17. Longchamp, P. F., C. Mauel, and D. Karamata. 1994. Lytic enzymes associated with defective prophages of *Bacillus subtilis*: sequencing and characterization of the region comprising the *N*-acetylmuramoyl-L-alanine amidase gene of prophage PBSX. *Microbiol.* **140**:1855-1867.
18. L  pez, R., E. Garc  a, P. Garc  a, and J. L. Garc  a. 1995. Architecture and domain interchange of the pneumococcal cell wall lytic enzymes. *Dev. Biol. Stand.* **85**:273-281.

19. Oda, Y., R. Nakayama, A. Kuroda, and J. Sekiguchi. 1993. Molecular cloning, sequence analysis, and characterization of a new cell wall hydrolase, CwlL, of *Bacillus licheniformis*. *Mol. Gen. Genet.* **241**:380-388.
20. Oshida, T., M. Sugai, H. Komatsuzawa, Y. M. Hong, H. Suginaka, and A. Tomasz. 1995. A *Staphylococcus aureus* autolysin that has an *N*-acetylmuramoyl-L-alanine amidase domain and an endo- β -*N*-acetylglucosaminidase domain: Cloning, sequence analysis, and characterization. *Proc. Natl. Acad. Sci. USA* **92**:285-289.
21. Pérez Martínez, G., J. Kok, G. Venema, J. M. van Dijk, H. Smith, and S. Bron. 1992. Protein export elements from *Lactococcus lactis*. *Mol. Gen. Genet.* **234**:401-411.
22. Potvin, C., D. Leclerc, G. Tremblay, A. Asselin, and G. Bellemare. 1988. Cloning, sequencing and expression of a *Bacillus* bacteriolytic enzyme in *Escherichia coli*. *Mol. Gen. Genet.* **214**:241-248.
23. Rashid, M. H., M. Mori, and J. Sekiguchi. 1995. Glucosaminidase of *Bacillus subtilis*: Cloning, regulation, primary structure and biochemical characterization. *Microbiol.* **141**:2391-2404.
24. Ruhland, G. J., M. Hellwig, G. Wanner, and F. Feldler. 1993. Cell-surface location of *Listeria*-specific protein p60: detection of *Listeria* cells by indirect immunofluorescence. *J. Gen. Microbiol.* **139**:609-616.
25. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
26. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463-5467.
27. Sanz, J. M., P. García, and J. L. García. 1996. Construction of a multifunctional pneumococcal murein hydrolase by module assembly. *Eur. J. Biochem.* **235**:601-605.
28. Sheehan, M. M., J. L. García, R. López, and P. García. 1996. Analysis of the catalytic domain of the lysin of the lactococcal bacteriophage Tuc2009 by chimeric gene assembling. *FEMS Microbiol. Lett.* **140**:23-28.
29. Smith, H., S. Eron, J. van Ee, and G. Venema. 1987. Construction and use of signal sequence selection vectors in *Escherichia coli* and *Bacillus subtilis*. *J. Bacteriol.* **169**:3321-3328.
30. Studier, F. W. and B. A. Moffatt. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189**:113-130.
31. Towbin, H., T. Staechelin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA* **76**:4350-4354.
32. van De Guchte, M., J. Kodde, J. M. B. M. van der Vossen, J. Kok, and G. Venema. 1990. Heterologous gene expression in *Lactococcus lactis* ssp. *lactis*: Synthesis, secretion, and processing of the *Bacillus subtilis* neutral protease. *Appl. Environ. Microbiol.* **56**:2606-2611.
33. van Dijk, J. M., A. de Jong, J. Vehmaanpera, G. Venema, and S. Bron. 1992. Signal peptidase I of *Bacillus subtilis*: patterns of conserved amino acids in prokaryotic and eukaryotic type I signal peptidases. *EMBO J.* **11**:2819-2828.
34. Vasala, A., M. Valkkila, J. Caldentey, and T. Alatossava. 1995. Genetic and biochemical characterization of the *Lactobacillus delbrueckii* subsp. *lactis* bacteriophage LL-H lysin. *Appl. Environ. Microbiol.* **61**:4004-4011.
35. Vielra, J. and J. Messing. 1991. New pUC-derived cloning vectors with different selectable markers and DNA replication origins. *Gene* **100**:189-194.
36. Wilson, R., R. Ainscough, K. Anderson, C. Baynes, M. Berks, J. Bonfield, J. Burton, M. Connell, T. Copsy, J. Cooper, and et al. 1994. 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*. *Nature* **368**:32-38.
- Zabarovsky, E. R. and G. Winberg. 1990. High efficiency electroporation of ligated DNA into bacteria. *Nucl. Acid Res.* **18**:5912.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: Rijksuniversiteit Groningen
- (B) STREET: Broerstraat 5
- (C) CITY: Groningen
- (D) STATE: Groningen
- (E) COUNTRY: the Netherlands
- (F) POSTAL CODE (ZIP): 9712 CP

(ii) TITLE OF INVENTION: Attaching substances to micro-organisms.

(iii) NUMBER OF SEQUENCES: 109

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(v) CURRENT APPLICATION DATA:

APPLICATION NUMBER: EP 97203539.8

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CGCGAATTC A GATTATGAAA CAATAAG
27

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

CGCGAATTCT TATGTCAGTA CAAGTTTTTG
30

10

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

15

(ii) MOLECULE TYPE: other nucleic acid

20

(iii) HYPOTHETICAL: NO

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

CGCGAATTCC TTATGAAGAA GCTCCGTC
28

30

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

35

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

CTTCAACAGA CAAGTCC
17

45

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

50

55

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

AGCAATACTA GTTTTATA
18

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

CGCGAATTCG CTAGCGTCGC TCAAATTCAA AGTGCG
36

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

AGGAGATCTG CGACTAACTC ATCAGAGG
28

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GGATCATGTA ACTCGCC
17

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GGAATCCATA TGCTTAATCA GTGAGG
26

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

GCATCCGTTG AAAGCGG
17

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs

EP 0 916 726 A1

(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GAATTCGATA TCTTTGAACA TAAATTG
27

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GATAAAATGA TTCCAAGC
18

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

CTCAAATTCA AAGTGCG
17

(2) INFORMATION FOR SEQ ID NO: 14:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

Tyr Thr Val Lys Ser Gly Asp Thr Leu Trp Gly Ile Ser
 Gln Arg Tyr
 1 5 10
 15
 Gly Ile Ser Val Ala Gln Ile Gln Ser Ala Asn Asn Leu
 Lys Ser Thr
 20 25
 30
 Ile Ile Tyr Ile Gly Gln Lys Leu Val Leu Thr
 35 40

(2) INFORMATION FOR SEQ ID NO: 15:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 45 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

Thr Thr Val Lys Val Lys Ser Gly Asp Thr Leu Trp Ala
 Leu Ser Val
 1 5 10
 15
 Lys Tyr Lys Thr Ser Ile Ala Gln Leu Lys Ser Trp Asn
 His Leu Ser
 20 25
 30
 Ser Asp Thr Ile Tyr Ile Gly Gln Asn Leu Ile Val Ser
 35 40 45

(2) INFORMATION FOR SEQ ID NO: 16:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 45 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Ser Ile His Lys Val Val Lys Gly Asp Thr Leu Trp Gly
 Leu Ser Val
 1 5 10
 15
 Lys Ser Gly Ser Pro Ile Ala Ser Ile Lys Ala Trp Asn
 His Leu Ser
 20 25
 30
 Ser Asp Thr Ile Leu Ile Gly Gln Tyr Leu Arg Ile Lys
 35 40 45

(2) INFORMATION FOR SEQ ID NO: 17:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 45 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(ix) FEATURE:

- (A) NAME/KEY: Protein
 (B) LOCATION:1..44
 (D) OTHER INFORMATION:/product= "OTHER"
 /label= X
 /note= "'X" stands for "unknown" amino

acid."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

Tyr Xaa Val Lys Xaa Gly Asp Thr Leu Xaa Xaa Ile Ala
 Xaa Xaa Xaa
 1 5 10
 15

Xaa Xaa Xaa Xaa Xaa Xaa Leu Xaa Xaa Xaa Asn Xaa Xaa
Leu Xaa Xaa

20

25

30

Xaa Xaa Xaa Ile Xaa Xaa Gly Gln Xaa Ile Xaa Val Xaa
35 40

(2) INFORMATION FOR SEQ ID NO: 18:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 45 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(ix) FEATURE:

- (A) NAME/KEY: Protein
- (B) LOCATION:1..44

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

His Xaa Ile Arg Xaa Xaa Glu Ser Val Xaa Xaa Leu Ser
Xaa Xaa Xaa
1 5 10

Xaa Xaa Xaa Xaa Xaa Xaa Ile Xaa Xaa Xaa Xaa Xaa Xaa
Ile Xaa Xaa
20 25

30

Xaa Xaa Xaa Leu Xaa Xaa Xaa Xaa Xaa Leu Xaa Xaa Ile
35 40

(2) INFORMATION FOR SEQ ID NO: 19:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 45 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(ix) FEATURE:

- (A) NAME/KEY: Protein
- (B) LOCATION:1..44

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

5 Xaa Xaa Leu Xaa Xaa Xaa Xaa Xaa Ile Xaa Xaa Val Xaa
Xaa Xaa Xaa 1 5 10
15
10 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
Xaa Xaa Xaa 20 25
30
15 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Val Xaa Leu Xaa
35 40

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 45 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

30 Thr Thr Tyr Thr Val Lys Ser Gly Asp Thr Leu Trp Gly
Ile Ser Gln 1 5 10
15
35 Arg Tyr Gly Ile Ser Val Ala Gln Ile Gln Ser Ala Asn
Asn Leu Lys 20 25
30
40 Ser Thr Ile Ile Tyr Ile Gly Gln Lys Leu Val Leu Thr
35 40 45

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 43 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

5 Val Lys Val Lys Ser Gly Asp Thr Leu Trp Ala Leu Ser
 Val Lys Tyr
 1 5 10
 15
 10 Lys Thr Ser Ile Ala Gln Leu Lys Ser Trp Asn His Leu
 Ser Ser Asp
 20 25
 30
 15 Thr Ile Tyr Ile Gly Gln Asn Leu Ile Val Ser
 35 40

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:
 20 (A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

25 (ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

His Lys Val Val Lys Gly Asp Thr Leu Trp Gly Leu Ser
 Val Lys Ser
 35 1 5 10
 15
 Gly Ser Pro Ile Ala Ser Ile Lys Ala Trp Asn His Leu
 Ser Ser Asp
 20 25
 40 30
 Thr Ile Leu Ile Gly Gln Tyr Leu Arg Ile Lys
 35 40

45 (2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
 50 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

55

(iii) HYPOTHETICAL: NO

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

10

Tyr Thr Val Lys Ser Gly Asp Thr Leu Asn Lys Ile Ala
Ala Gln Tyr
1 5 10
15

15

Gly Val Ser Val Ala Asn Leu Arg Ser Trp Asn Gly Ile
Ser Gly Asp
20 25
30

20

Leu Ile Phe Val Gly Gln Lys Leu Ile Val Lys
35 40

(2) INFORMATION FOR SEQ ID NO: 24:

25

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

35

Tyr Thr Val Lys Ser Gly Asp Thr Leu Asn Lys Ile Ala
Ala Gln Tyr
1 5 10
15

40

Gly Val Thr Val Ala Asn Leu Arg Ser Trp Asn Gly Ile
Ser Gly Asp
20 25
30

45

Leu Ile Phe Val Gly Gln Lys Leu Ile Val Lys
35 40

(2) INFORMATION FOR SEQ ID NO: 25:

50

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

55

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

Tyr Thr Ile Lys Ser Gly Asp Thr Leu Asn Lys Ile Ala
Ala Gln Tyr 1 5 10
15
Gly Val Ser Val Ala Asn Leu Arg Ser Trp Asn Gly Ile
Ser Gly Asp 20 25
30
Leu Ile Phe Ala Gly Gln Lys Ile Ile Val Lys
35 40

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

His Thr Val Lys Ser Gly Asp Ser Leu Trp Gly Leu Ser
Met Gln Tyr 1 5 10
15
Gly Ile Ser Ile Gln Lys Ile Lys Gln Leu Asn Gly Leu
Ser Gly Asp 20 25
30
Thr Ile Tyr Ile Gly Gln Thr Leu Lys Val Gly
35 40

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 amino acids
(B) TYPE: amino acid

(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein
(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

Tyr Thr Val Lys Ser Gly Asp Ser Val Trp Gly Ile Ser
His Ser Phe
1 5 10
15
Gly Ile Thr Met Ala Gln Leu Ile Glu Trp Asn Asn Ile
Lys Asn Asn
20 25
30
Phe Ile Tyr Pro Gly Gln Lys Leu Thr Ile Lys
35 40

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein
(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

Tyr Thr Val Lys Ser Gly Asp Ser Val Trp Lys Ile Ala
Asn Asp His
1 5 10
15
Gly Ile Ser Met Asn Gln Leu Ile Glu Trp Asn Asn Ile
Lys Asn Asn
20 25
30
Phe Val Tyr Pro Gly Gln Gln Leu Val Val Ser
35 40

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 43 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

Tyr Thr Val Lys Ala Gly Glu Ser Val Trp Ser Val Ser
Asn Lys Phe
1 5 10
15

Gly Ile Ser Met Asn Gln Leu Ile Gln Trp Asn Asn Ile
Lys Asn Asn
20 25
30

Phe Ile Tyr Pro Gly Gln Lys Leu Ile Val Lys
35 40

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

Tyr Thr Val Lys Ala Gly Glu Ser Val Trp Gly Val Ala
Asn Lys Asn
1 5 10
15

Gly Ile Ser Met Asn Gln Leu Ile Glu Trp Asn Asn Ile
Lys Asn Asn
20 25
30

Phe Ile Tyr Pro Gly Gln Lys Leu Ile Val Lys
35 40

(2) INFORMATION FOR SEQ ID NO: 31:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

Tyr Thr Val Lys Ala Gly Glu Ser Val Trp Gly Val Ala
 Asn Lys His
 1 5 10
 15
 His Ile Thr Met Asp Gln Leu Ile Glu Trp Asn Asn Ile
 Lys Asn Asn
 20 25
 30
 Phe Ile Tyr Pro Gly Gln Glu Val Ile Val Lys
 35 40

(2) INFORMATION FOR SEQ ID NO: 32:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

Tyr Thr Val Lys Ala Gly Glu Ser Val Trp Gly Val Ala
 Asp Ser His
 1 5 10
 15
 Gly Ile Thr Met Asn Gln Leu Ile Glu Trp Asn Asn Ile
 Lys Asn Asn
 20 25
 30
 Phe Ile Tyr Pro Gly Gln Gln Leu Ile Val Lys
 35 40

(2) INFORMATION FOR SEQ ID NO: 33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

Tyr Val Val Lys Gln Gly Asp Thr Leu Ser Gly Ile Ala
 Ser Asn Trp

1 5 10
 15

Gly Thr Asn Trp Gln Glu Leu Ala Arg Gln Asn Ser Leu
 Ser Asn Pro

20 25
 30

Asn Met Ile Tyr Ala Gly Gln Val Ile Ser Phe Thr
 35 40

(2) INFORMATION FOR SEQ ID NO: 34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

Tyr Thr Val Gln Ser Gly Asp Asn Leu Ser Ser Ile Ala
 Ile Leu Leu

1 5 10
 15

Gly Thr Thr Val Gln Ser Leu Val Ser Met Asn Gly Ile
 Ser Asn Pro

20 25
 30

Asn Leu Ile Tyr Ala Gly Gln Thr Leu Asn Tyr
35 40

(2) INFORMATION FOR SEQ ID NO: 35:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

Tyr Ile Val Lys Gln Gly Asp Thr Leu Ser Gly Ile Ala
Ser Asn Leu
1 5 10
15

Gly Thr Asn Trp Gln Glu Leu Ala Arg Gln Asn Ser Leu
Ser Asn Pro
20 25
30

Asn Met Ile Tyr Ser Gly Gln Val Ile Ser Leu Thr
35 40

(2) INFORMATION FOR SEQ ID NO: 36:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

Tyr Thr Val Gln Ser Gly Asp Asn Leu Ser Ser Ile Ala
Arg Arg Leu
1 5 10
15

Gly Thr Thr Val Gln Ser Leu Val Ser Met Asn Gly Ile
Ser Asn Pro

10 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

Gly Val Thr Val Ala Gln Leu Gln Glu Trp Asn Asn Ile
 Glu Asp Pro

20 25

30 30

35 (2) INFORMATION FOR SEQ ID NO: 38:

(ii) MOLECULE TYPE: protein
(iii) HYPOTHETICAL: NO

50 Tyr Lys Val Lys Ser Gly Asp Asn Leu Thr Lys Ile Ala
Lys Lys His 5 10
15

Asn Thr Thr Val Ala Thr Leu Leu Lys Leu Asn Pro Ser
Ile Lys Asp

30 20 25

Pro Asn Met Ile Arg Val Gly Gln Thr Ile Asn Val Thr
35 40 45

(2) INFORMATION FOR SEQ ID NO: 39:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 45 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

His Lys Val Lys Ser Gly Asp Thr Leu Ser Lys Ile Ala
Val Asp Asn

1 5 10

Lys Thr Thr Val Ser Arg Leu Met Ser Leu Asn Pro Glu
Ile Thr Asn

20 25

Pro Asn His Ile Lys Val Gly Gln Thr Ile Arg Leu Ser
35 40 45

(2) INFORMATION FOR SEQ ID NO: 40:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 45 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

His Val Val Lys Lys Gly Asp Thr Leu Ser Glu Ile Ala
Lys Lys Ile

EP 0 916 726 A1

15 1 5 10
 5 Lys Thr Ser Thr Lys Thr Leu Leu Glu Leu Asn Pro Thr
 Ile Lys Asn 20 25
 30
 10 Pro Asn Lys Ile Tyr Val Gly Gln Arg Ile Asn Val Gly
 35 40 45

(2) INFORMATION FOR SEQ ID NO: 41:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 44 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

Tyr Thr Val Val Ser Gly Asp Ser Trp Trp Lys Ile Ala
 Gln Arg Asn 1 5 10
 15
 Gly Leu Ser Met Tyr Thr Leu Ala Ser Gln Asn Gly Lys
 Ser Ile Tyr 20 25
 30
 Ser Thr Ile Tyr Pro Gly Asn Lys Leu Ile Ile Lys
 35 40

(2) INFORMATION FOR SEQ ID NO: 42:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 44 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

Tyr Thr Val Val Ser Gly Asp Ser Trp Trp Lys Ile Ala
 Gln Arg Asn
 1 5 10
 5 15

Gly Leu Ser Met Tyr Thr Leu Ala Ser Gln Asn Gly Lys
 Ser Ile Tyr
 20 25
 10 30

Ser Thr Ile Tyr Pro Gly Asn Lys Leu Ile Ile Lys
 35 40

(2) INFORMATION FOR SEQ ID NO: 43:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 44 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

Ile Leu Ile Arg Pro Gly Asp Ser Leu Trp Tyr Phe Ser
 Asp Leu Phe
 1 5 10
 15

Lys Ile Pro Leu Gln Leu Leu Leu Asp Ser Asn Arg Asn
 Ile Asn Pro
 20 25
 30

Gln Leu Leu Gln Val Gly Gln Arg Ile Gln Ile Pro
 35 40

(2) INFORMATION FOR SEQ ID NO: 44:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 44 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

5 Tyr Thr Ile Thr Gln Gly Asp Ser Leu Trp Gln Ile Ala
 Gln Asn Lys 1 5 10
 15
 10 Asn Leu Pro Leu Asn Ala Ile Leu Leu Val Asn Pro Glu
 Ile Gln Pro 20 25
 30
 15 Ser Arg Leu His Ile Gly Gln Thr Ile Gln Val Pro
 35 40

(2) INFORMATION FOR SEQ ID NO: 45:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

30 His Ile Val Lys Lys Gly Glu Ser Leu Gly Ser Leu Ser
 Asn Lys Tyr 1 5 10
 15
 35 His Val Lys Val Ser Asp Ile Ile Lys Leu Asn Gln Leu
 Lys Arg Lys 20 25
 30
 40 Thr Leu Trp Leu Asn Glu Ser Ile Lys Ile Pro
 35 40

(2) INFORMATION FOR SEQ ID NO: 46:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 45 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

5 His Lys Val Thr Asn Lys Asn Gln Thr Leu Tyr Ala Ile
 Ser Arg Glu
 1 5 10
 15
 Tyr Asn Ile Pro Val Asn Ile Leu Leu Ser Leu Asn Pro
 10 His Leu Lys
 20 25
 30
 Asn Gly Lys Val Ile Thr Gly Gln Lys Ile Lys Leu Arg
 35 40 45
 15

(2) INFORMATION FOR SEQ ID NO: 47:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

30 Val Val Val Glu Ala Gly Asp Thr Leu Trp Gly Ile Ala
 Gln Ser Lys
 1 5 10
 15
 Gly Thr Thr Val Asp Ala Ile Lys Lys Ala Asn Asn Leu
 35 Thr Thr Asp
 20 25
 30
 Lys Ile Val Pro Gly Gln Lys Leu Gln Val Asn
 35 40
 40

(2) INFORMATION FOR SEQ ID NO: 48:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

5 His Ala Val Lys Ser Gly Asp Thr Ile Trp Ala Leu Ser
Val Lys Tyr
 1 5 10
15
 Gly Val Ser Val Gln Asp Ile Met Ser Trp Asn Asn Leu
10 Ser Ser Ser
 20 25
30
 Ser Ile Tyr Val Gly Gln Lys Leu Ala Ile Lys
15 35 40

(2) INFORMATION FOR SEQ ID NO: 49:

(i) SEQUENCE CHARACTERISTICS:
20 (A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

25 (iii) HYPOTHETICAL: NO

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

 His Asn Val Lys Ser Gly Asp Thr Ile Trp Ala Leu Ser
Val Lys Tyr
 1 5 10
35 15
 Gly Val Ser Val Gln Asp Ile Met Ser Trp Asn Asn Leu
 Ser Ser Ser
 20 25
40 30
 Ser Ile Tyr Val Gly Gln Lys Pro Ala Ile Lys
 35 40

45 (2) INFORMATION FOR SEQ ID NO: 50:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
50 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:

10

Val Val Val Glu Ala Gly Asp Thr Leu Trp Gly Ile Ala
Gln Asp Lys
1 5 10
15

15

Gly Thr Thr Val Asp Ala Leu Lys Lys Ala Asn Asn Leu
Thr Ser Asp
20 25
30

Lys Ile Val Pro Gly Gln Lys Leu Gln Ile Thr
35 40

20

(2) INFORMATION FOR SEQ ID NO: 51:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

25

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:

35

Tyr Thr Val Lys Ser Gly Asp Thr Ile Trp Ala Leu Ser
Ser Lys Tyr
1 5 10
15

40

Gly Thr Ser Val Gln Asn Ile Met Ser Trp Asn Asn Leu
Ser Ser Ser
20 25
30

45

Ser Ile Tyr Val Gly Gln Val Leu Ala Val Lys
35 40

(2) INFORMATION FOR SEQ ID NO: 52:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

50

55

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 52:

10

Tyr Thr Val Lys Ser Gly Asp Thr Leu Ser Lys Ile Ala
Thr Thr Phe

1 5 10
15

15

Gly Thr Thr Val Ser Lys Ile Lys Ala Leu Asn Gly Leu
Asn Ser Asp

20 25
30

20

Asn Leu Gln Val Gly Gln Val Leu Lys Val Lys
35 40

(2) INFORMATION FOR SEQ ID NO: 53:

(i) SEQUENCE CHARACTERISTICS:

25

- (A) LENGTH: 43 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

30

(iii) HYPOTHETICAL: NO

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:

Val Val Val Glu Ala Gly Asp Thr Leu Trp Gly Ile Ala
Gln Asp Asn

1 5 10
15

40

Gly Thr Thr Val Asp Ala Leu Lys Lys Ala Asn Lys Leu
Thr Thr Asp

20 25
30

45

Lys Ile Val Pro Gly Gln Lys Leu Gln Val Thr
35 40

(2) INFORMATION FOR SEQ ID NO: 54:

50

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 amino acids
- (B) TYPE: amino acid

55

(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:

Val	His	Thr	Val	Lys	Ser	Gly	Asp	Thr	Ile	Trp	Ala	Leu	Ser
Lys	Tyr												
1				5						10			
15													
	Gly	Ala	Ser	Val	Gln	Asp	Leu	Met	Ser	Trp	Asn	Asn	Leu
Ser	Ser	Ser											
				20						25			
30													
	Ser	Ile	Tyr	Val	Gly	Gln	Asn	Ile	Ala	Val	Lys		
				35					40				

(2) INFORMATION FOR SEQ ID NO: 55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:

Ser	Tyr	Thr	Val	Lys	Ser	Gly	Asp	Thr	Leu	Gly	Lys	Ile	Ala
Thr	Phe												
1				5						10			
15													
	Gly	Thr	Thr	Val	Ser	Lys	Ile	Lys	Ala	Leu	Asn	Gly	Leu
Thr	Ser	Asp											
				20						25			
30													
	Asn	Leu	Gln	Val	Gly	Asp	Val	Leu	Lys	Val	Lys		
				35					40				

(2) INFORMATION FOR SEQ ID NO: 56:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 43 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 56:

Val Val Val Glu Ala Gly Asp Thr Leu Trp Gly Ile Ala
Gln Ser Lys
1 5 10
15
Gly Thr Thr Val Asp Ala Leu Lys Lys Ala Asn Asn Leu
Thr Ser Asp
20 25
30
Lys Ile Val Pro Gly Gln Lys Leu Gln Val Thr
35 40

(2) INFORMATION FOR SEQ ID NO: 57:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:

His Thr Val Lys Ser Gly Asp Thr Ile Trp Ala Leu Ser
Val Lys Tyr
1 5 10
15
Gly Ala Ser Val Gln Asp Leu Met Ser Trp Asn Asn Leu
Ser Ser Ser
20 25
30
Ser Ile Tyr Val Gly Gln Lys Ile Ala Val Lys
35 40

(2) INFORMATION FOR SEQ ID NO: 58:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:

Tyr Thr Val Lys Ser Gly Asp Ser Leu Ser Lys Ile Ala
 Asn Thr Phe
 1 5 10
 15
 Gly Thr Ser Val Ser Lys Ile Lys Ala Leu Asn Asn Leu
 Thr Ser Asp
 20 25
 30
 Asn Leu Gln Val Gly Thr Val Leu Lys Val Lys
 35 40
 25

(2) INFORMATION FOR SEQ ID NO: 59:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 59:

Val Val Val Ala Ser Gly Asp Thr Leu Trp Gly Ile Ala
 Ser Lys Thr
 1 5 10
 15
 Gly Thr Thr Val Asp Gln Leu Lys Gln Leu Asn Lys Leu
 Asp Ser Asp
 20 25
 30
 Arg Ile Val Pro Gly Gln Lys Leu Thr Ile Lys
 35 40
 50

(2) INFORMATION FOR SEQ ID NO: 60:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60:

Tyr Lys Val Lys Ser Gly Asp Thr Ile Trp Ala Leu Ser
 Val Lys Tyr
 1 5 10
 15
 Gly Val Pro Val Gln Lys Leu Ile Glu Trp Asn Asn Leu
 Ser Ser Ser
 20 25
 30
 Ser Ile Tyr Val Gly Gln Thr Ile Ala Val Lys
 35 40

(2) INFORMATION FOR SEQ ID NO: 61:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:

Tyr Lys Val Gln Asn Gly Asp Ser Leu Gly Lys Ile Ala
 Ser Leu Phe
 1 5 10
 15
 Lys Val Ser Val Ala Asp Leu Thr Asn Trp Asn Asn Leu
 Asn Ala Thr
 20 25
 30

Ile Thr Ile Tyr Ala Gly Gln Glu Leu Ser Val Lys
35 40

(2) INFORMATION FOR SEQ ID NO: 62:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 62:

Tyr Val Val Ser Thr Gly Asp Thr Leu Ser Ser Ile Leu
 Asn Gln Tyr 5 10

Gly Ile Asp Met Gly Asp Ile Ser Gln Leu Ala Ala Ala
 Asp Lys Glu 20 25
 30

Leu Arg Asn Leu Lys Ile Gly Gln Gln Leu Ser Trp Thr
35 40 45

(2) INFORMATION FOR SEQ ID NO: 63:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 45 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 63:

Tyr Thr Val Thr Glu Gly Asp Thr Leu Lys Asp Val Leu
Val Leu Ser 1 5 10

Gly Leu Asp Asp Ser Ser Val Gln Pro Leu Ile Ala Leu
Asp Pro Glu

Leu Ala His Leu Lys Ala Gly Gln Gln Phe Tyr Trp Ile
35 40 45

10 (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

Tyr Thr Val Lys Lys Gly Asp Thr Leu Phe Tyr Ile Ala
Trp Ile Thr
1 5 10

Gly Asn Asp Phe Arg Asp Leu Ala Gln Arg Asn Asn Ile
Gln Ala Pro

Tyr Ala Leu Asn Val Gly Gln Thr Leu Gln Val Gly

 35 40

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(iii) HYPOTHETICAL: NO

50 Tyr Lys Val Asn Lys Gly Asp Thr Met Phe Leu Ile Ala
Tyr Leu Ala
1 5 10

15

Gly Ile Asp Val Lys Glu Leu Ala Ala Leu Asn Asn Leu
Ser Glu Pro

20

25

5 30

Asn Tyr Asn Leu Ser Leu Gly Gln Val Leu Lys Ile Ser
35 40 45

(2) INFORMATION FOR SEQ ID NO: 66:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 44 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 66:

Tyr Lys Val Arg Lys Gly Asp Thr Met Phe Leu Ile Ala
Tyr Ile Ser

1

5

10

15

Gly Met Asp Ile Lys Glu Leu Ala Thr Leu Asn Asn Met
Ser Glu Pro

20

25

30 30

Tyr His Leu Ser Ile Gly Gln Val Leu Lys Ile Ala
35 40

(2) INFORMATION FOR SEQ ID NO: 67:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 44 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 67:

Tyr Ile Val Arg Arg Gly Asp Thr Leu Tyr Ser Ile Ala
Phe Arg Phe

15 1 5 10
 5 Gly Trp Asp Trp Lys Ala Leu Ala Ala Arg Asn Gly Ile
 Ala Pro Pro 20 25
 30
 10 Tyr Thr Ile Gln Val Gly Gln Ala Ile Gln Phe Gly
 35 40

(2) INFORMATION FOR SEQ ID NO: 68:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 44 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 68:

 His Gln Val Lys Glu Gly Glu Ser Leu Trp Gln Ile Ser
 Gln Ala Phe
 1 5 10
 15
 30 Gln Val Asp Ala Lys Ala Ile Ala Leu Ala Asn Ser Ile
 Ser Thr Asp 20 25
 30
 35 Thr Glu Leu Gln Ala Gly Gln Val Leu Asn Ile Pro
 35 40

(2) INFORMATION FOR SEQ ID NO: 69:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 44 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 69:

Ile Met Val Arg Gln Gly Asp Thr Val Thr Val Leu Ala
 Arg Arg Phe
 1 5 10
 15

Gly Val Pro Glu Lys Glu Ile Leu Lys Ala Asn Gly Leu
 Lys Ser Ala
 20 25
 30

Ser Gln Val Glu Pro Gly Gln Arg Leu Val Ile Pro
 35 40

(2) INFORMATION FOR SEQ ID NO: 70:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 70:

Tyr Thr Val Arg Ser Gly Asp Thr Leu Ser Ser Ile Ala
 Ser Arg Leu
 1 5 10
 15

Gly Val Ser Thr Lys Asp Leu Gln Gln Trp Asn Lys Leu
 Arg Gly Ser
 20 25
 30

Lys Leu Lys Pro Gly Gln Ser Leu Thr Ile Gly
 35 40

(2) INFORMATION FOR SEQ ID NO: 71:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 42 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 71:

5 Tyr Arg Val Arg Lys Gly Asp Thr Leu Ser Ser Ile Ala
 Lys Arg His
 1 5 10
 15
 10 Gly Val Asn Ile Lys Asp Val Met Arg Trp Asn Ser Asp
 Thr Ala Asn
 20 25
 30
 15 Leu Gln Pro Gly Asp Lys Leu Thr Leu Phe
 35 40

(2) INFORMATION FOR SEQ ID NO: 72:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 44 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 72:

30 His Val Val Lys Pro Gly Asp Thr Val Asn Asp Ile Ala
 Lys Ala Asn
 1 5 10
 15
 35 Gly Thr Thr Ala Asp Lys Ile Ala Ala Asp Asn Lys Leu
 Ala Asp Lys
 20 25
 30
 40 Asn Met Ile Lys Pro Gly Gln Glu Leu Val Val Asp
 35 40

(2) INFORMATION FOR SEQ ID NO: 73:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 73:

Ile Lys Val Lys Lys Gly Asp Thr Leu Trp Asp Leu Ser
 Arg Lys Tyr
 1 5 10
 15
 Asp Thr Thr Ile Ser Lys Ile Lys Ser Glu Asn His Leu
 Arg Ser Asp
 20 25
 30
 Ile Ile Tyr Val Gly Gln Thr Leu Ser Ile Asn
 35 40

(2) INFORMATION FOR SEQ ID NO: 74:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74:

Tyr Lys Val Lys Ser Gly Asp Ser Leu Trp Lys Ile Ser
 Lys Lys Tyr
 1 5 10
 15
 Gly Met Thr Ile Asn Glu Leu Lys Lys Leu Asn Gly Leu
 Lys Ser Asp
 20 25
 30
 Leu Leu Arg Val Gly Gln Val Leu Lys Leu Lys
 35 40

(2) INFORMATION FOR SEQ ID NO: 75:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 75:

5 Tyr Lys Val Lys Ser Gly Asp Ser Leu Ser Lys Ile Ala
 Ser Lys Tyr 1 5 10
 15
 10 Gly Thr Thr Val Ser Lys Leu Lys Ser Leu Asn Gly Leu
 Lys Ser Asp 20 25
 30
 15 Val Ile Tyr Val Asn Gln Val Leu Lys Val Lys
 35 40

(2) INFORMATION FOR SEQ ID NO: 76:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 44 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 76:

30 Cys Ile Val Gln Gln Glu Asp Thr Ile Glu Arg Leu Cys
 Glu Arg Tyr 1 5 10
 35 15
 Glu Ile Thr Ser Gln Gln Leu Ile Arg Met Asn Ser Leu
 Ala Leu Asp 20 25
 40 30
 Asp Glu Leu Lys Ala Gly Gln Ile Leu Tyr Ile Pro
 35 40

(2) INFORMATION FOR SEQ ID NO: 77:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 44 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 77:

10

Tyr Thr Val Lys Arg Gly Asp Thr Leu Tyr Arg Ile Ser
Arg Thr Thr
1 5 10
15

15

Gly Thr Ser Val Lys Glu Leu Ala Arg Leu Asn Gly Ile
Ser Pro Pro
20 25
30

Tyr Thr Ile Glu Val Gly Gln Lys Leu Lys Leu Gly
35 40

20

(2) INFORMATION FOR SEQ ID NO: 78:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

25

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 78:

35

His Val Val Lys Ala Gly Glu Thr Ile Asp Ser Ile Ala
Ala Gln Tyr
1 5 10
15

40

Gln Leu Val Pro Ala Thr Leu Ile Ser Val Asn Asn Gln
Leu Ser Ser
20 25
30

45

Gly Gln Val Thr Pro Gly Gln Thr Ile Leu Ile Pro
35 40

(2) INFORMATION FOR SEQ ID NO: 79:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

50

55

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 79:

10

Met Val Lys Gln Gly Asp Thr Leu Ser Ala Ile Ala Ser
Gln Tyr Arg
1 5 10
15

15

Thr Thr Thr Asn Asp Ile Thr Glu Thr Asn Glu Ile Pro
Asn Pro Asp
20 25
30

20

Ser Leu Val Val Gly Gln Thr Ile Val Ile Pro
35 40

(2) INFORMATION FOR SEQ ID NO: 80:

(i) SEQUENCE CHARACTERISTICS:

25

- (A) LENGTH: 44 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

30

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 80:

35

Tyr Asp Val Lys Arg Gly Asp Thr Leu Thr Ser Ile Ala
Arg Gln Phe
5 10
15

40

Asn Thr Thr Ala Ala Glu Leu Ala Arg Val Asn Arg Ile
Gln Leu Asn
20 25
30

45

Thr Val Leu Gln Ile Gly Phe Arg Leu Tyr Ile Pro
35 40

(2) INFORMATION FOR SEQ ID NO: 81:

(i) SEQUENCE CHARACTERISTICS:

50

- (A) LENGTH: 43 amino acids
- (B) TYPE: amino acid

55

(C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 81:

Ile Lys Val Lys Ser Gly Asp Ser Leu Trp Lys Leu Ala
Gln Thr Tyr

1 5 10
15

Asn Thr Ser Val Ala Ala Leu Thr Ser Ala Asn His Leu
Ser Thr Thr

20 25
30

Val Leu Ser Ile Gly Gln Thr Leu Thr Ile Pro
35 40

(2) INFORMATION FOR SEQ ID NO: 82:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 43 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 82:

Tyr Thr Val Lys Ser Gly Asp Ser Leu Trp Leu Ile Ala
Asn Glu Phe

1 5 10
15

Lys Met Thr Val Gln Glu Leu Lys Lys Leu Asn Gly Leu
Ser Ser Asp

20 25
30

Leu Ile Arg Ala Gly Gln Lys Leu Lys Val Ser
35 40

(2) INFORMATION FOR SEQ ID NO: 83:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 83:

Tyr Lys Val Gln Leu Gly Asp Ser Leu Trp Lys Ile Ala
 Asn Lys Val
 1 5 10
 15
 Asn Met Ser Ile Ala Glu Leu Lys Val Leu Asn Asn Leu
 Lys Ser Asp
 20 25
 30
 Thr Ile Tyr Val Asn Gln Val Leu Lys Thr Lys
 35 40

(2) INFORMATION FOR SEQ ID NO: 84:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 84:

Tyr Thr Val Lys Ser Gly Asp Ser Leu Trp Lys Ile Ala
 Asn Asn Tyr
 1 5 10
 15
 Asn Leu Thr Val Gln Gln Ile Arg Asn Ile Asn Asn Leu
 Lys Ser Asp
 20 25
 30
 Val Leu Tyr Val Gly Gln Val Leu Lys Leu Thr
 35 40

(2) INFORMATION FOR SEQ ID NO: 85:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 85:

Tyr Thr Val Lys Ser Gly Asp Ser Leu Trp Val Ile Ala
 Gln Lys Phe
 1 5 10
 15
 Asn Val Thr Ala Gln Gln Ile Arg Glu Lys Asn Asn Leu
 Lys Thr Asp
 20 25
 30
 Val Leu Gly Val Gly Gln Lys Leu Val Ile Ser
 35 40

(2) INFORMATION FOR SEQ ID NO: 86:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 44 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 86:

Thr Glu Ile Lys Ser Gly Asp Ser Cys Trp Asn Ile Ala
 Ser Asn Ala
 1 5 10
 15
 Lys Ile Ser Val Glu Arg Leu Gln Gln Leu Asn Lys Gly
 Met Lys Cys
 20 25
 30
 Asp Lys Leu Pro Leu Gly Asp Lys Leu Cys Leu Ala
 35 40

(2) INFORMATION FOR SEQ ID NO: 87:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 87:

	Tyr	Thr	Ile	Lys	Ser	Gly	Asp	Thr	Leu	Asn	Lys	Ile	Ser
Ala	Gln	Phe											
	1			5					10				
	Gly	Val	Ser	Val	Ala	Asn	Leu	Arg	Ser	Trp	Asn	Gly	Ile
Lys	Gly	Asp											
			20					25					
	Leu	Ile	Phe	Ala	Gly	Gln	Thr	Ile	Ile	Val	Lys		
			35					40					

(2) INFORMATION FOR SEQ ID NO: 88:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 44 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 88:

	Leu	Lys	Leu	Lys	Ala	Glu	Asp	Thr	Cys	Pro	Lys	Ile	Trp
Ser	Ser	Gln											
	1			5					10				
	Lys	Leu	Ser	Glu	Arg	Gln	Phe	Leu	Gly	Met	Asn	Glu	Gly
Met	Asp	Cys											
			20					25					

EP 0 916 726 A1

Asp Lys Leu Lys Val Gly Lys Glu Val Cys Val Ala
35 40

(2) INFORMATION FOR SEQ ID NO: 89:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 89:

His Lys Ile Gln Lys Gly Asp Thr Cys Phe Lys Ile Trp
Thr Thr Asn
1 5 10

Lys Ile Ser Glu Lys Gln Phe Arg Asn Leu Asn Lys Gly
Leu Asp Cys
20 25
30

Asp Lys Leu Glu Ile Gly Lys Glu Val Cys Ile Ser
35 40

(2) INFORMATION FOR SEQ ID NO: 90:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 90:

Leu Lys Ile Lys Glu Gly Asp Thr Cys Tyr Asn Ile Trp
Thr Ser Gln
1 5 10

Lys Ile Ser Glu Gln Glu Phe Met Glu Leu Asn Lys Gly
Leu Asp Cys

EP 0 916 726 A1

30
 5 Asp Lys Leu Glu Ile Gly Lys Glu Val Cys Val Thr
 35 40

(2) INFORMATION FOR SEQ ID NO: 91:

10 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 44 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

15 (ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 91:

Tyr Arg Phe Lys Lys Gly Asp Thr Cys Tyr Lys Ile Trp
 Thr Ser His
 1 5 10

25 15 Lys Met Ser Glu Lys Gln Phe Arg Ala Leu Asn Arg Gly
 Ile Asp Cys
 20 25

30 30 Asp Arg Leu Val Pro Gly Lys Glu Leu Cys Val Gly
 35 40

(2) INFORMATION FOR SEQ ID NO: 92:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 44 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

40 (ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 92:

Ile Thr Val Lys Pro Gly Asp Thr Cys Phe Ser Ile Trp
 Thr Ser Gln
 1 5 10
 15

55

Lys Met Thr Gln Gln Gln Phe Met Asp Ile Asn Pro Glu
Leu Asp Cys

20

25

30

Asp Lys Leu Glu Ile Gly Lys Glu Val Cys Val Thr
35 40

(2) INFORMATION FOR SEQ ID NO: 93:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 44 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 93:

Val Lys Ile Asn Pro Gly Asp Thr Cys Phe Asn Ile Trp
Thr Ser Gln
1 5 10

15

Arg Met Thr Gln Gln Gln Phe Met Asp Leu Asn Lys Arg
Leu Asp Cys

20

25

30

Asp Lys Leu Glu Val Gly Lys Glu Val Cys Val Thr
35 40

(2) INFORMATION FOR SEQ ID NO: 94:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 44 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 94:

Val Gln Ile Asn Pro Gly Asp Thr Cys Phe Lys Ile Trp
Ser Ala Gln

EP 0 916 726 A1

15 1 5 10
 5 Lys Leu Thr Glu Gln Gln Phe Met Glu Leu Asn Lys Gly
 Leu Asp Cys 20 25
 30
 10 Asp Arg Leu Glu Val Gly Lys Glu Val Cys Ile Ala
 35 40

(2) INFORMATION FOR SEQ ID NO: 95:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 44 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 95:

Thr Glu Val Lys Glu Gly Asp Thr Cys Phe Lys Ile Trp
 Ser Ala His
 1 5 10
 15
 Lys Ile Thr Glu Gln Gln Phe Met Glu Met Asn Arg Gly
 Leu Asp Cys 20 25
 30
 35 Asn Arg Leu Glu Val Gly Lys Glu Val Cys Ile Val
 35 40

(2) INFORMATION FOR SEQ ID NO: 96:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 44 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 96:

Ile Lys Val Lys Glu Gly Asp Thr Cys Phe Lys Ile Trp
Ser Ala Gln

1 5 10

Lys Met Thr Glu Gln Gln Phe Met Glu Met Asn Arg Gly
Leu Asp Cys

20 25

Asn Lys Leu Met Val Gly Lys Glu Val Cys Val Ser
35 40

(2) INFORMATION FOR SEQ ID NO: 97:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 97:

Ala Thr Ile Thr Pro Gly Asn Thr Cys Phe Asn Ile Ser
Val Ala Tyr

1 5 10

Gly Ile Asn Leu Thr Asp Leu Gln Lys Thr Tyr Asp Cys
Lys Ala Leu

20 25

Glu Val Gly Asp Thr Ile Cys Val Ser
35 40

(2) INFORMATION FOR SEQ ID NO: 98:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 98:

5 Ile Glu Val Ile Lys Gly Asp Thr Cys Trp Phe Leu Glu
 Asn Ala Phe
 1 5 10
 15
 Lys Thr Asn Gln Thr Glu Met Glu Arg Ala Asn Glu Gly
 10 Val Lys Cys
 20 25
 30
 Asp Asn Leu Pro Ile Gly Arg Met Met Cys Val Trp
 35 40
 15

(2) INFORMATION FOR SEQ ID NO: 99:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 44 amino acids
 (B) TYPE: amino acid
 20 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

25 (iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 99:

30 His Thr Ile Lys Ser Gly Asp Thr Cys Trp Lys Ile Ala
 Ser Glu Ala
 1 5 10
 15
 Ser Ile Ser Val Gln Glu Leu Glu Gly Leu Asn Ser Lys
 35 Lys Ser Cys
 20 25
 30
 Ala Asn Leu Ala Val Gly Leu Ser Glu Gln Glu Phe
 35 40
 40

(2) INFORMATION FOR SEQ ID NO: 100:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 44 amino acids
 (B) TYPE: amino acid
 45 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

50 (ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 100:

Ile His Val Lys Glu Gly Asp Thr Cys Tyr Thr Ile Trp
 Thr Ser Gln
 1 5 10
 15
 His Leu Thr Glu Lys Gln Phe Met Asp Met Asn Glu Glu
 Leu Asn Cys
 20 25
 30
 Gly Met Leu Glu Ile Gly Asn Glu Val Cys Val Asp
 35 40

(2) INFORMATION FOR SEQ ID NO: 101:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 41 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 101:

Ala Thr Val Thr Pro Gly Ser Ser Cys Tyr Thr Ile Ser
 Ala Ser Tyr
 1 5 10
 15
 Gly Leu Asn Leu Ala Glu Leu Gln Thr Thr Tyr Asn Cys
 Asp Ala Leu
 20 25
 30
 Gln Val Asp Asp Thr Ile Cys Val Ser
 35 40

(2) INFORMATION FOR SEQ ID NO: 102:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 44 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 102:

Ile Glu Ile Leu Asn Gly Asp Thr Cys Gly Phe Leu Glu
 Asn Ala Phe
 1 5 10
 15
 Gln Thr Asn Asn Thr Glu Met Glu Ile Ala Asn Glu Gly
 Val Lys Cys
 10 20 25
 30
 Asp Asn Leu Pro Ile Gly Arg Met Met Cys Val Trp
 35 40
 15

(2) INFORMATION FOR SEQ ID NO: 103:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 45 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 103:

His Thr Val Gln Lys Lys Glu Thr Leu Tyr Arg Ile Ser
 Met Lys Tyr
 1 5 10
 15
 Tyr Lys Ser Arg Thr Gly Glu Glu Lys Ile Arg Ala Tyr
 Asn His Leu
 20 25
 30
 Asn Gly Asn Asp Val Tyr Thr Gly Val Leu Asp Ile Pro
 35 40 45
 40

(2) INFORMATION FOR SEQ ID NO: 104:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 49 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 104:

10

Tyr Thr Leu Lys Thr Gly Glu Ser Val Ala Gln Leu Ser
Lys Ser Gln
1 5 10
15

15

Gly Ile Ser Val Pro Val Ile Trp Ser Leu Asn Lys His
Leu Tyr Ser
20 25
30

20

Ser Glu Ser Glu Met Met Lys Ala Ser Pro Gly Gln Gln
Ile Ile Leu
35 40 45
Pro

(2) INFORMATION FOR SEQ ID NO: 105:

25

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 49 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

30

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 105:

40

Tyr Thr Leu Lys Thr Gly Glu Thr Val Ala Asp Leu Ser
Lys Ser Gln
1 5 10
15

45

Asp Ile Asn Leu Ser Thr Ile Trp Ser Leu Asn Lys His
Leu Tyr Ser
20 25
30

50

Ser Glu Ser Glu Met Met Lys Ala Ala Pro Gly Gln Gln
Ile Ile Leu
35 40 45
Pro

55

(2) INFORMATION FOR SEQ ID NO: 106:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 51 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 106:

Ile	Glu	Val	Gln	Gln	Gly	Asp	Thr	Leu	Trp	Ser	Ile	Ala
Asp	Gln	Val										
1					5					10		
15												
Ala	Asp	Thr	Lys	Lys	Ile	Asn	Lys	Asn	Asp	Phe	Ile	Glu
Trp	Val	Ala										
				20				25				
25												
30												
Asp	Lys	Asn	Gln	Leu	Gln	Thr	Ser	Asp	Ile	Gln	Pro	Gly
Asp	Glu	Leu										
			35				40			45		
30												
Val	Ile	Pro										
		50										

(2) INFORMATION FOR SEQ ID NO: 107:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 55 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 107:

Tyr	Thr	Val	Lys	Tyr	Gly	Asp	Thr	Leu	Ser	Thr	Ile	Ala
Glu	Ala	Met										
1					5					10		
15												

EP 0 916 726 A1

Gly Ile Asp Val His Val Leu Gly Asp Ile Asn His Ile
Ala Asn Ile 20 25

Asp Leu Ile Phe Pro Asp Thr Ile Leu Thr Ala Asn Tyr
Asn Gln His 35 40 45

Gly Gln Ala Thr Thr Leu Thr
50 55

(2) INFORMATION FOR SEQ ID NO: 108:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 58 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 108:

Tyr Thr Val Lys Lys Gly Asp Thr Leu Trp Asp Ile Ala
Gly Arg Phe 1 5 10
15

Tyr Gly Asn Ser Thr Gln Trp Arg Lys Ile Trp Asn Ala
Asn Lys Thr 20 25
30

Ala Met Ile Lys Arg Ser Lys Arg Asn Ile Arg Gln Pro
Gly His Trp 35 40 45

Ile Phe Pro Gly Gln Lys Leu Lys Ile Pro
50 55

(2) INFORMATION FOR SEQ ID NO: 109:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 58 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 109:

5 Tyr Thr Val Lys Lys Gly Asp Thr Leu Trp Asp Leu Ala
 Gly Lys Phe
 1 5 10
 15

10 Tyr Gly Asp Ser Thr Lys Trp Arg Lys Ile Trp Lys Val
 Asn Lys Lys
 20 25
 30

15 Ala Met Ile Lys Arg Ser Lys Arg Asn Ile Arg Gln Pro
 Gly His Trp
 35 40 45

20 Ile Phe Pro Gly Gln Lys Leu Lys Ile Pro
 50 55

25

Claims

1. A proteinaceous substance comprising at least one stretch of amino acids derived from a first micro-organism which substance is capable of attaching to a cell-wall of a second micro-organism, said stretch of amino acids having a sequence corresponding to a consensus sequence listed in figure 10.
2. A substance according to claim 1 wherein said stretch of amino acids has a sequence corresponding to a sequence selected from those listed in figure 11, or a homologous sequence derived from another species.
3. A substance according to claim 1 or 2 wherein said second micro-organism is a non-recombinant micro-organism.
4. A substance according to claim 3 wherein said second micro-organism is selected from any of the group of Gram-positive bacteria and Gram-negative bacteria.
5. A substance according to any one of claims 1 to 4 additionally comprising a reactive group.
6. A substance according to claim 5 wherein said reactive group is selected from the group of antibiotics, hormones, aromatic substances and reporter molecules.
7. A substance according to claim 5 wherein said reactive group is selected from the group of antigenic determinants, enzymes, (single-chain) antibodies or fragments thereof, polyhistidyl tags, fluorescing proteins, binding proteins or peptides.
8. A nucleic acid molecule encoding a proteinaceous substance according to claim 7.
9. A vector comprising a nucleic acid molecule according to claim 8.
10. A micro-organism or expression system comprising a nucleic acid molecule according to claim 8 or a vector according to claim 9 or expressing a substance according to any one of claims 1 to 5 or 7.
11. A method for attaching a substance to the cell wall of a micro-organism comprising the use of an attaching peptide which comprises a stretch of amino acids having a sequence corresponding to the consensus amino acid

sequence provided in figure 10.

12. A method for attaching a substance to the cell wall of a micro-organism comprising the use of an attaching peptide derived from any one of the proteins listed in figure 11.

13. A method according to claim 11 or 12 comprising the use of an attaching peptide which is derived from the major peptidoglycan hydrolase of *Lactococcus lactis*.

14. Use of a method according to any one of claims 11 to 13 to attach a substance to the cell-wall of a micro-organism.

15. A micro-organism to which a substance has been attached using a method according to any one of claims 11 to 13.

16. A pharmaceutical composition comprising a substance according to any of claims 1 to 7 or a micro-organism according to claim 10 or 15.

17. A vaccine comprising a substance according to any of claims 1 to 7 or a micro-organism according to claim 10 or 15.

18. A foodstuff comprising a substance according to any of claims 1 to 7 or a micro-organism according to claim 10 or 15.

19. A process for the preparation of a foodstuff comprising the use of a substance according to any of claims 1 to 7 or a micro-organism according to claim 10 or 15.

20. Use of a substance according to any of claims 1 to 7 or a micro-organism according to claim 10 or 15 in a diagnostic test.

21. Use of a substance according to any of claims 1 to 7 or a micro-organism according to claim 10 or 15 in a bioadsorption process.

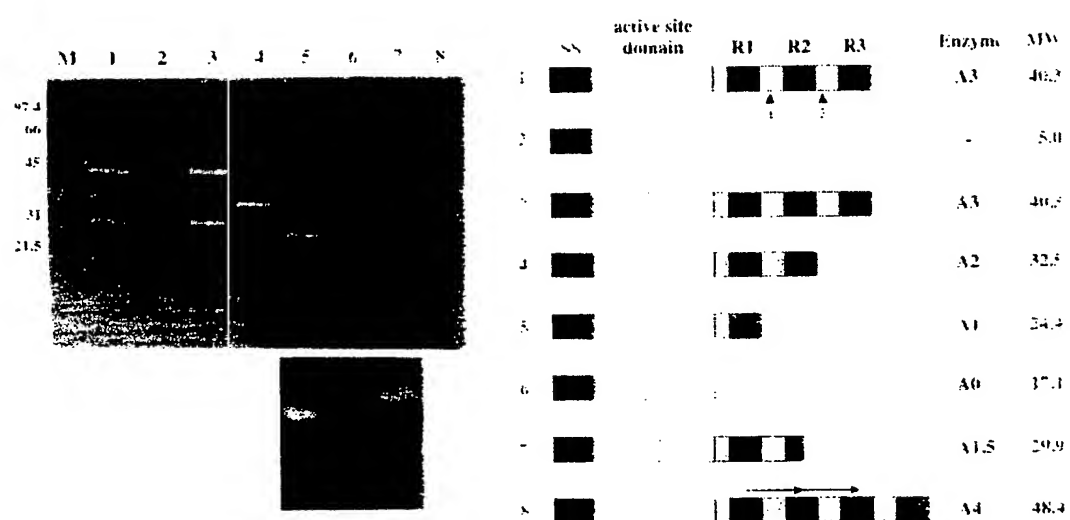


FIG. 1

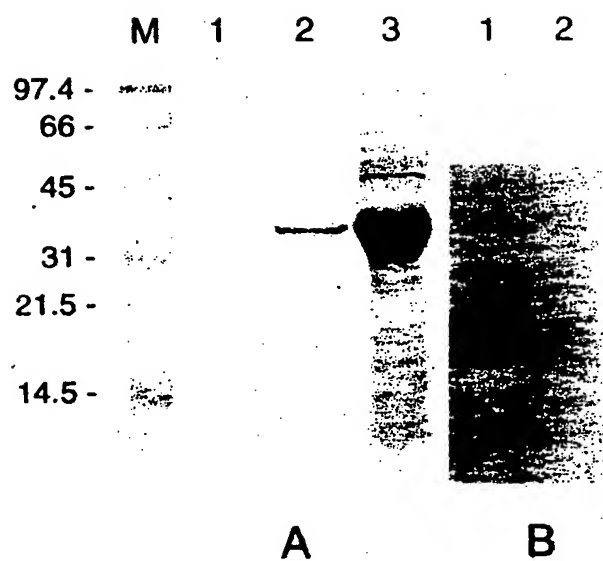


FIG. 2

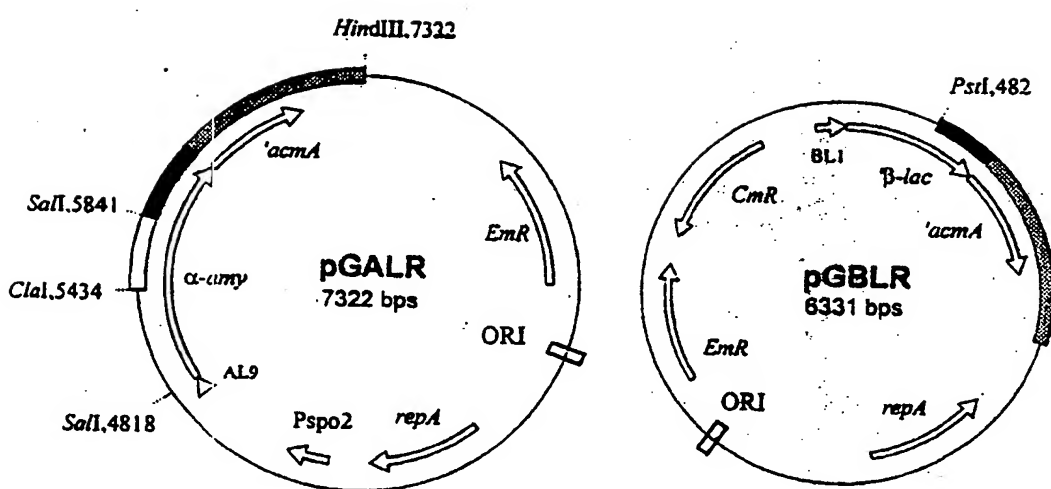


Fig. 3

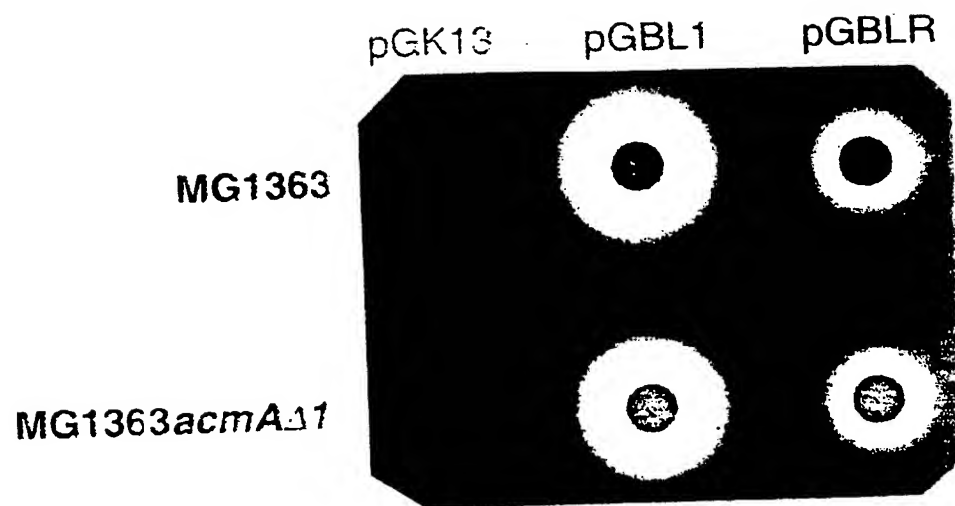


FIG. 4

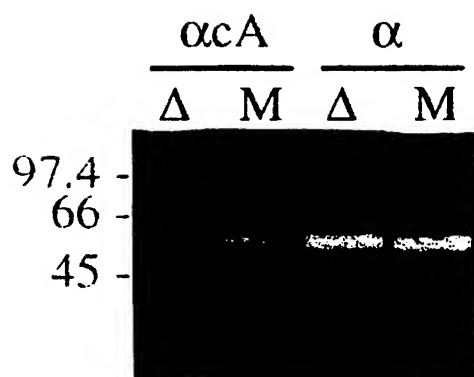


FIG. 5

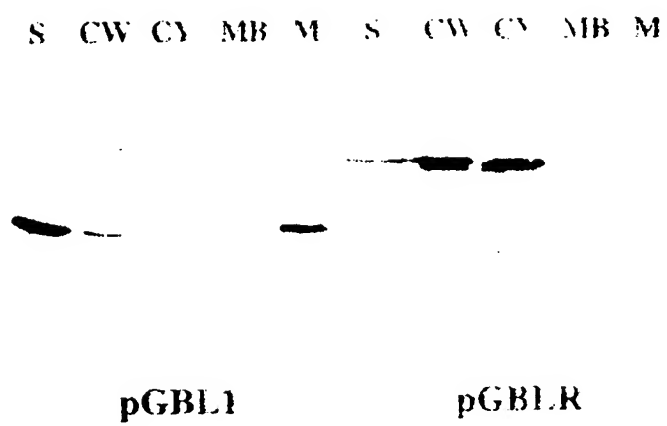


FIG. 6

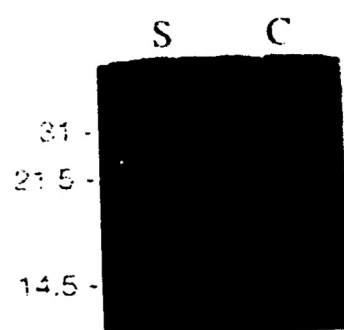


FIG. 7

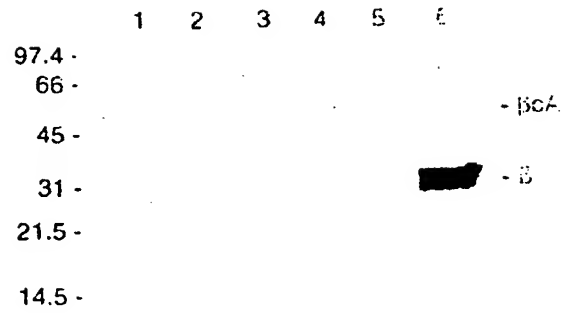
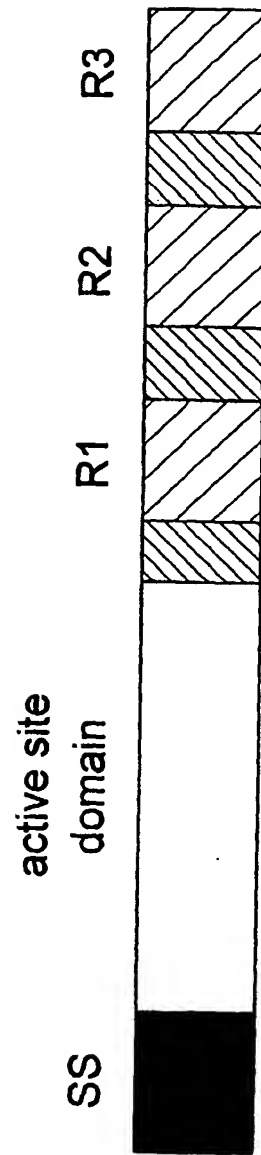


FIG. 8



1

FIG. 9

Consensus:

FIG. 10

FIGURE 11

1	<i>Lactococcus lactis</i>	* <i>acmA</i>	YTVKSGDTLWGISQRYGISVAQIQSAN NLKST IYIGQKLVLT VKVKSQDTLWALS VKYKTSIAQLKSWN HLSSD TIYIGQNLIYS HKVVKGDTLWGLSKQSGSPIASIKAWN HLSSD TILIGQYLRIK YTVKSGDTLNKIAAQYGVSVANLRSWN GISGD LIFVGQKLIVK YTVKSGDTLNKIAAQYGVTVANLRSWN GISGD LIFVGQKLIVK YTVKSGDTLNKIAAQYGVSVANLRSWN GISGD LIFAGQKIIVK YTVKSGDTLNKISAQFGVSVANLRSWN GIKGD LIFAGQTIIVK HTVKSQDSLWGLSMQYGISIQIKIKQLN GLSGD TIYIGQTLKVG YTVKSGDSVWGISHSFGITMAQLIEWN NIKNN FIYPGQKLTIK YTVKSGDSVWKIANDHGISMNQLIEWN NIKNN FVYPGQQLVVS YTVKAGESVWSVSNKFGISMNQLIQWN NIKNN FIYPGQKLIVK YTVKAGESVWGVANKNGISMNQLIEWN NIKNN FIYPGQKLIVK YTVKAGESVWGVANKHHITMDQLIEWN NIKNN FIYPGQEVIVK YTVKAGESVWGVADSHGITMNQLIEWN NIKNN FIYPGQQLIVK YVVKQGDTLTSGIASNWGTNWQELARQN SLSPNPMIYAGQVISFT YTVQSGDNLSSIAILLGTTVQSLVSMN GISNPNLIYAGQTLNY YIVKQGDTLTSGIASNLGTNWQELARQN SLSPNPMIYSGQVISLT YTVQSGDNLSSIAARLGTTVQSLVSMN GISNPNLIYAGQTLNY YVVKQGDTLTSGIARAFGTVTAQLQEWN NIEDPNLIRVGQVLIVS YVVKSGDNLTAKAKHNTTVA TLKLNPSIKDPNMIRVGQTLINVT HKVKSQDTLSKIAVDNKTTSRLMSLNPEITPNPNHIKVGQTLRLS HVVKKGDTLSEIAKKIKTSTKTLLLELNPTIKNPNKIYVGQRINVG YKIKRGETLTGIAKKNKTTVSQLMKNPNIKNANNIYAGQTLRLK YTVVSGDSWWMKIAQRNGLSMYTLASQN GKSIYSTIYPGNKLIK ILIRPGDSLWYFSDLFKIPQLQLLLDSNRNINPQ LLQVGQRIQIP YTVITQGDSLWQIAQKNLPLNAILLVNPEIQPS RLHIGQTIQVP HIVKKGESLGLSNKYHVKSVDIIKLN QLKRK TLWLNESIKIP HKVTKNQTLTYAISREYNIPVNILLSLPHLKNK KVTGQKIKLR VVVEAGDTLWQIAQSKGTTVDAIKKAN NLTTD KIVPGQKLQVN HAVKSGDTIWL SVKYGVSVQDIMSWN NLSSS SIYVGQKLAIK VVVEAGDTLWQIAQSKGTTVDAIKKAN NLTTD KIVPGQKLQVN HNKSGDTIWL SVKYGVSVQDIMSWN NLSSS SIYVGQKPAIK
2	<i>Enterococcus faecalis</i>	* <i>autolysin</i>	
3	<i>Enterococcus hirae</i>	* <i>mur2</i>	
4	<i>Lactococcus</i> ØTuc2009	* <i>lys</i>	
5	<i>Lactococcus</i> ØLC3	* <i>lysB</i>	
6	<i>Bacillus</i> ØPBSX	* <i>xylA</i>	
7	<i>Bacillus</i> ØPZA (=Ø29)	* <i>orf15</i>	
8	<i>Bacillus</i> ØB103	* <i>orf15</i>	
9	<i>Bacillus</i> Øgle	* <i>lys</i>	
10	<i>Bacillus sphaericus</i>	* <i>pepI</i>	
11	<i>Haemophilus influenzae</i>	* <i>amiB</i>	
12	<i>Listeria monocytogenes</i>	* <i>p60</i>	
13	<i>Listeria innocua</i>	* <i>p60</i>	

14	<i>Listeria ivanovii</i>	* P60	VVEAGDTLWGIAQDKGTTVDALKKAN NLTS D KIVPGQKLQIT YTVKSGDTIWALSSKYGTSVQNMISWN NLSS S SIYVGQVLAVK YTVKSGDTLSKIATTFGTTVSKIKALN GLNS D NLQVGQVLKVK VVEAGDTLWGIAQDNNGTTVDALKKAN KLTT D KIVPGQKLQVT HTVKSGDTIWALSVKYGASVQDLMISWN NLSS S SIYVGQNIQAVK YTVKSGDTLKGIASTFGTTVSKIKALN GLTS D NLQVGDVLKVK VVEAGDTLWGIAQSKGTTVDALKKAN NLTS D KIVPGQKLQVT HTVKSGDTIWALSVKYGASVQDLMISWN NLSS S SIYVGQKIQAVK YTVKSGDSLKIANTFGTTSVSKIKALN NLTS D NLQVGTVLKVK VVASGDTLWGIAASKTGTVDQLKQLN KLDSD RIVPGQKLTIK YVKSGDTIWALSVKYGVFQVKLIEWN NLSS S SIYVGQTLQAVK YKVQNGDSLGIASLKFVSVADLTNWN NLNATITTIYAGQELSVK YVVGSTGDTLSSILNQYIGIDMGDISQLAAADKELRNKIGQQLSWT YTVTEGDTLKDVLVLSGLDDSSVQPLIALDPELAHLKAGQQFYWI YTVKKGDTLFIYIAWITGNDFRDIAQRN NIQAPYALNVGQTLQVG YKVNKGDTMFLIAYLAGIDVKELAAALNNSEPNVNLGLQVVKIS YKVRKGDTMFLIAYISGMDIKELATLN NMSEPYHLSIGQVLKIA YIVRRGDTLYSIAFRFGWDWKALAAARN GIAPPYTIQVGQAIQFG HQVKEGESLWQISQAFQVDKAKAIALAN SISTDTLQAGQVNLNIP IMVRQGDTVTVLARRFGVPEKEILKAN GLKSASQVEPGQRLVIP YTVRS GDTLSSIASRLGVSTKDLOQWN KLRS S KLKPGQSLTIG YVRKKGDTLSSIAKRHG VNIKDVMRWN SDTA NLQPGDKLTLF HVKPGD TVNDIAKANGTTADKIAADN KLADKNMIKPGQELVVD IKVKKGDTLWDL SRKYDTTISKIKSEN HLRSD IIVVGQTL SIN YKVKSGDSLWKISKKYGMTINELKKN GLKSD LLRVGQVLKVK YKVKSGDSL SKIASKYGTTVSKLKS LN GLKSD VIYVNQVLKVK CIVQQEDTIERLCERYEITSQQILRMN SLALDDELKAGQILYIP YTVKR GDTLYRISR TTGTSVKELARLN GISPPYTTIEVGQKLKLG HVKAGETIDSI AAQYQLVPATLISVN NQLSSGQVTPGQTLIP MVKQGD TLSAIASQYRTTNDITETN EIPNPD SLVVGQTLVIP YDVKR GDTLTSIARQFN TTAELARVN RIQLNTVLQIGFRLYIP
15	<i>Listeria seeligeri</i>	* P60	
16	<i>Listeria welshimeri</i>	* P60	
17	<i>Listeria grayi</i>	* P60	
18	<i>Escherichia coli</i>	* yeba	
19	<i>Haemophilus influenzae</i>	* yeba	
20	<i>Escherichia coli</i>	nlpD	
21	<i>Haemophilus influenzae</i>	lppB	
22	<i>Haemophilus somnus</i>	lppB	
23	<i>Pseudomonas aeruginosa</i>	lppB	
24	<i>Synechocystis</i>	nlpD	
25	<i>Sinohizobium meliloti</i>	nlpD	
26	<i>Escherichia coli</i>	dniR	
27	<i>Staphylococcus aureus</i>	Prota	
28	<i>Bacillus subtilis</i>	papQ	
29	<i>Bacillus subtilis</i>	spoVID	
30	<i>Escherichia coli</i>		
31	<i>Synechocystis</i> sp.		
32	<i>Bacillus subtilis</i>	yaaH	

33	<i>Bacillus subtilis</i>	yhdD	IKVKSGLSWKLAQTYNTSVAALTSAN HLSTT VLSIGQTLTIP YTVKSGDSLWLIANEFKMTVQELKKLN GLSSD LIRAGQKLKVS YKVQLGDSLWKIANKVNMSIAELKVLN NLKSD TIYVNQVLKTK YTVKSGDSLWKIANNNYNTVQQIRNIN NLKSD VLYVGQVLKLT YTVKSGDSLWVIAQKFNVTAAQIREKN NLKTD VLGVGQKLVIS
34	<i>Caenorhabditis elegans</i>		TEIKSGDCWNIAASNAKISVERLQQLN KGMKCDKLPGLGDKLCLA LKLKAEDTCFKIWSQKLSERQFLGMN EGMDCDKLVGKEVCVA HKIQGDTCFKIWTNNKISEKQFRNLN KGLDCDKLEIGKEVCIS LKIKEGDTCYNIWTSQKISEQEFMELN KGLDCDKLEIGKEVCVT YRFKKGDTCYKIWTSHKMSEKQFRALN RGIDCDRLVPGKELCVG ITVKPGDTCFSIWT SQMTQQQFMDIN PELDCDKLEIGKEVCVT VKINPGDTCFNIWTSQMTQQQFMDLN KRLDCDKLEVGEVCVT VQINPGDTCFKIWSAQKLEQQFMEFN KGLDCDRLEVGEVCIA TEVKEGDTCFKIWSAHKITEQQFMEFN RGLDCNRLEVGEVCIV IKVKEGDTCFKIWSAQKMTQQFMEFN RGLDCNKL MVGKEVCVS ATTPGNTCFNISVAYGINLT DLQ KTYDCKALEVGDTICVS IEVIKGDTCWFLENAFQTNNTMEIAN EGVKCDNLPIGRMVCVW
35	<i>Caenorhabditis elegans</i>		HTIKSGDTCWKIAEASISVQLEGLN SKKSCANLAVGLSEQEF IHVKEGDTCTYTIWTSQHLTEKQFMDMN EELNCGMLEIGNEVCVD ATVTPGSSCYTISASYGLNLAELQTTY NCDALQVDDTICVS IEILNGDTCGFLENAFQTNNTMEIAN EGVKCDNLPIGRMVCVW
36	<i>Bacillus subtilis</i>		HTVQKKTLYRISMYYKSRTGEEKIRAYNHNGNDVYTGQVLDIP
37	<i>Citrobacter freundii</i>	# eae	YTLKTGESVAQLSKSQGISVPVIWSLNKHLYSSESEMMKASPGQIILP
38	<i>Escherichia coli</i>	# eae	YTLKTGETVADLSKSDINLSTIWSLNKHLYSSESEMMKAAPGQIILP
39	<i>Bacillus subtilis</i>	# yneA	IEVQQGDTLWSIADQVADTKKINKNDFIEWVADKNLQTSDIQPGDELVIP
40	<i>Streptococcus pyogenes</i>	#	YTVKYGDTLSTIAEAMGIDVHVLGDINHIANIDLIFPDTILTANYNHGQATTLT
41	<i>Bacillus subtilis</i>	# yqbp	YTVKKGDTLWDIAGRFYGNSTQWRKIWNANKTAMIKRSKRNIQPGHWIFPGQKLKIP
42	<i>Bacillus subtilis</i>	#	YTVKKGDTLWDLAGKFYGDSTKWRKIWKVNKKAMIKRSKRNIQPGHWIFPGQKLKIP

1	245-287 (33) 321-363 (31) 395-437	437	U17696	muramidase
2	363-405 (25) 431-473 (25) 499-541 (25) 567-609 (19) 629-671	671	P37710	muramidase
3	257-299 (38) 338-380 (33) 414-456 (32) 489-531 (33) 565-607 (15) 623-665	666	P39046	muramidase
4	332-375 (10) 386-428	428	L31364	glycosidase (muramidase)
5	333-376 (10) 387-429	429	U04309	muramidase
6	161-204	297	P39800	N-acetylmuramoyl-L-alanine amidase
7	163-207 (6) 214-258	258	P11187 P07540	lysozyme
8	165-209 (9) 219-263	263	X99260	lysozyme
9	399-442	442	1926389	lysin
10	3- 46 (6) 53- 96	396	X69507	carboxypeptidase I
11	294-336 387-430	432	P44493	N-acetylmuramoyl-L-alanine amidase
12	30- 72 (130) 203-245	478	M80351	adherence and invasion protein P60
13	30- 72 (130) 201-243	481	Q01836	adherence and invasion protein P60

14	30- 72(125) 198-240(73) 314-356	524	Q01837	adherence and invasion protein P60
15	30- 72(127) 200-242(75) 320-362	523	Q01838	adherence and invasion protein P60
16	30- 72(127) 198-240(75) 316-358	524	M80348	adherence and invasion protein P60
17	30- 72(104) 177-219(79) 299-342	511	Q01835	adherence and invasion protein P60
18	77-121	419	P24204	homologous to endopeptidase of <i>Staphylococcus</i>
19	131-174	475	P44693	homologous to endopeptidase of <i>Staphylococcus</i>
20	123-166	379	P33648	Lipoprotein
21	147-190	405	P44833	Lipoprotein
22	120-164	279	L10653	Lipoprotein
23	69-113	297	P45682	Lipoprotein
24	87-130	715	D90915	Lipoprotein
25	166-209	512	U81296	Lipoprotein
26	113-155(16) 172-213	222	P23931	involved in hexaheme nitrite reductase expression
27	431-474	524	A04512	protein A
28	28- 70(17) 88-130(20)	334	U38819	phosphatase-associated protein, P60 homolog
29	151-193	575	P37963	stage VI sporulation protein D
30	525-568	259	U28375	Hypothetical protein
31	50- 93	245	D90907	Hypothetical protein
32	4- 47 1- 43(5) 49- 92	427	p37531	Hypothetical protein

33	29- 71 94-136 176-218 242-284 309-353	488	2226145	Hypothetical protein
34	23- 66 78-121 143-186 208-251 271-314 335-378 402-445 467-510 548-591 636-679 746-786 795-838	1614	U64836	PUTATIVE ENDOCHITINASE
35	23- 66 118-161 187-226 236-279	1484	U70858	PUTATIVE ENDOCHITINASE
36	191-136	240	L47648	
37	65-113	936	Q07591	
38	65-113	934	P43261	
39	40- 90	105	Z73234	
40	47-103	339	U09352	
41	177-234	235	G1225954	
42	161-218	219	P45932	



European Patent
Office

EUROPEAN SEARCH REPORT

Application Number
EP 97 20 3539

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
Y	CHEMICAL ABSTRACTS, vol. 127, no. 3, 21 July 1997 Columbus, Ohio, US; abstract no. 31403, PIARD, J.-C. ET AL: "Cell wall anchoring of the Streptococcus pyogenes M6 protein in various lactic acid bacteria" XP002063307 * abstract * & J. BACTERIOL. (1997), 179(9), 3068-3072 CODEN: JOBAAY;ISSN: 0021-9193, ---	1	C12N15/31 C12N1/20 C07K14/315 C07K14/195 C07K14/37 C12N9/36 A61K38/02 A23L1/03 G01N33/68 B01J20/00
Y	US 5 616 686 A (FISCHETTI VINCENT A ET AL) * claims * ---	1	
Y	CHEMICAL ABSTRACTS, vol. 112, no. 7, 12 February 1990 Columbus, Ohio, US; abstract no. 51896, PANCHOLI, VIJAYKUMAR ET AL: "Identification of an endogenous membrane anchor-cleaving enzyme for group A streptococcal M protein. Its implication for the attachment of surface proteins in Gram-positive bacteria" XP002063308 * abstract * & J. EXP. MED. (1989), 170(6), 2119-33 CODEN: JEMEAV;ISSN: 0022-1007, --- -/--	1	TECHNICAL FIELDS SEARCHED (Int.Cl.6) C07K C12N
The present search report has been drawn up for all claims			
Place of search THE HAGUE		Date of completion of the search 24 April 1998	Examiner Delanghe, L
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons ----- & : member of the same patent family, corresponding document	

EPO FORM 1503 03.82 (P04C01)



European Patent
Office

EUROPEAN SEARCH REPORT

Application Number
EP 97 20 3539

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
Y	CHEMICAL ABSTRACTS, vol. 106, no. 13, 30 March 1987 Columbus, Ohio, US; abstract no. 98247, TAYLOR, M. J. ET AL: "Is there evidence for a common amino acid sequence of proteins with membrane attaching ability?" XP002063309 * abstract * & J. THEOR. BIOL. (1986), 121(3), 293-306 CODEN: JTBIAP;ISSN: 0022-5193, ---	1	TECHNICAL FIELDS SEARCHED (Int.Cl.6)
Y	CHEMICAL ABSTRACTS, vol. 123, no. 9, 28 August 1995 Columbus, Ohio, US; abstract no. 106450, BUIST, GIRBE ET AL: "Molecular cloning and nucleotide sequence of the gene encoding the major peptidoglycan hydrolase of Lactococcus lactis, a muramidase needed for cell separation" XP002063310 * abstract * & J. BACTERIOL. (1995), 177(6), 1554-63 CODEN: JOBAAY;ISSN: 0021-9193, ---	1	
Y	WO 95 31561 A (QUEST INT ;BUIST GIRBE (NL); VENEMA GERARD (NL); KOK JAN (NL); LED) * claims * ---	1	
Y	WO 95 09232 A (ALSTYNE DIANE VAN ;SHARMA LAWRENCE RAJENDRA (CA)) * claims * -----	1	
The present search report has been drawn up for all claims			
Place of search THE HAGUE		Date of completion of the search 24 April 1998	Examiner Delanghe, L
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document			

EPO FORM 1503 03/92 (P4/C01)